ANALYSIS II

CONTINUITY AND DIFFERENTIABILITY

DR JANET DYSON

HILARY TERM 2013

Draft 2nd January 2013

Contents

Message from the lecturer

Acknowledgement

These lectures have been developed by a number of lecturers over the years. I would particularly like to thank Professor Roger Heath-Brown who gave the first lectures for this course in its present form in 2003 and Dr Brian Stewart and Dr Zhongmin Qian who allowed me to adapt their lecture notes and use their LaTeX files.

Lectures

To get the most out of the course you must attend the lectures. There will be more explanation in the lectures than there is in the notes.

On the other hand I will not put everything on the board which is in the printed notes. In some places I have put in extra examples which I will not have time to demonstrate in the lectures. There is also some extra material (which is generally in a smaller font) which I have put in for interest but which I do not regard as central to the course.

Numbering system:

In the printed notes there are 16 sections. Within each section there are subsections. Theorems, definitions, etc are numbered consecutively within each section. So for example Theorem 1.3 is the third result in Section 1. I will use the numbering in the printed notes, even though I will omit some subsections in the lectures, so the numbering will no longer be consecutive.

Exercise sheets

The weekly problem sheets which accompany the lectures are an integral part of the course. In Analysis above all you will only understand the definitions and theorems by using them.

I assume that week 1 tutorials are being devoted to the final sheets from the Michaelmas Term courses.

I suggest that the problem sheets for this course are tackled in tutorials in weeks 2–8, with the 8th sheet used a vacation work for a tutorial in the first week of Trinity Term.

Corrections

Please email any corrections to me at `Janet.Dyson@mansfield.ox.ac.uk`

Notation:

I will use this notation (which was used in courses on MT) throughout.

- $\mathbb{C}$: set of all complex numbers - the complex plane.
- $\mathbb{R}$: set of all real numbers - the real line; $\mathbb{R} \subset \mathbb{C}$.
- $\mathbb{Q}$: the rational numbers; $\mathbb{Q} \subset \mathbb{R}$.
- $\mathbb{N}$: the natural numbers, $1, 2, \ldots$; $\mathbb{N} \subset \mathbb{Q}$.
- $\forall$: "for all" or"for every" or "whenever".
- $\exists$: "there exist(s)" or "there is (are)".
- Sometimes I will write "s. t." for "such that" , "resp." for "respectively", "iff" for "if and only if".

Recall the following definition from 'Introduction to Pure Mathematics' last term

If $a, b \in \mathbb{R}$ then we define intervals as follows:

$$\begin{aligned}
(a, b) &:= \{x \in \mathbb{R} : a < x < b\} \\
[a, b] &:= \{x \in \mathbb{R} : a \leq x \leq b\} \\
(-\infty, a) &:= \{x \in \mathbb{R} : x < a\},
\end{aligned}$$

etc.

1 Limits of Functions

1.1 Sequence limits and completeness

This course builds on the ideas from Analysis I and also uses many of the results from that course. I have put some of the most important results from Analysis I in these notes but I will not write them on the board in the lecture. However, I will begin by recalling the definition of limits for *sequences*.

Definition 1.1. *A sequence (z_n) of real (or complex) numbers has **limit** l, if* $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$*, such that* $\forall n > N$

$$|z_n - l| < \varepsilon.$$

We denote this by '$z_n \to l$ as $n \to \infty$' or by '$\lim_{n\to\infty} z_n = l$'.

Definition 1.2. *A sequence (z_n) of real (or complex) numbers **converges** if it has a limit l.*

Often we prove things by contradiction. We start by assuming that what we want is not true. That means we have to be able to write down the **contrapositive** of a proposition. We can do this mechanically: working from the left change every $\forall$ into $\exists$, every $\exists$ into $\forall$ and negate the simple proposition at the end.

For example, by the first definition, a sequence (z_n) *does not* converge to l [1], if and only if $\exists \varepsilon > 0$, such that $\forall k \in \mathbb{N}$, $\exists n_k > k$ such that

$$|z_{n_k} - l| \geqslant \varepsilon.$$

Definition 1.3. (z_n) *is called a **Cauchy sequence** if* $\forall \varepsilon > 0$ $\exists N \in \mathbb{N}$ *such that* $\forall n, m > N$

$$|z_n - z_m| < \varepsilon.$$

Here is the key theorem, sometimes called **The General Principle for Convergence**:

Theorem (Cauchy's Criterion). *A sequence (z_n) of real (or complex) numbers converges if and only if it is a Cauchy sequence.*

When mathematicians say that the real number system $\mathbb{R}$ and the complex number system $\mathbb{C}$ are *complete* what they mean is that this theorem is true. There are no sequences which look as though they converge but don't, there are no 'gaps' in the real number line.

[1] i.e. either (z_n) diverges, or $z_n \to a \neq l$

According to Cauchy's criterion, (z_n) diverges [i.e. has no finite limit], if and only if $\exists\varepsilon > 0$, such that $\forall k \in \mathbb{N}$, there exist [at least] two integers n_{k_1}, $n_{k_2} > k$ s. t. $|z_{n_{k_1}} - z_{n_{k_2}}| \geqslant \varepsilon$.

The following theorem demonstrates the "compactness" of a bounded subset.

Theorem. *(**The Bolzano–Weierstrass Theorem**) Any bounded sequence in $\mathbb{R}$ (or in $\mathbb{C}$) has a subsequence which converges to a point in $\mathbb{R}$ (in $\mathbb{C}$).*

1.2 Limit points

We want to define what is meant by the limit of a function. Intuitively f has a limit l at the point p if the values of $f(x)$ are close to l when x is close to (but not equal to) p. But for the definition of limit to be meaningful it is necessary that f is defined at 'enough' points close to p. So we are interested only in points p that x can get close to. Such points are limit points of the domain of f.

Definition 1.4. *Let $E \subseteq \mathbb{R}$ (or $\mathbb{C}$). A point $p \in \mathbb{R}$ (or $\mathbb{C}$) is called a **limit point** (or an **accumulation point**) of E, if $\forall\varepsilon > 0$, there exists $z \in E$ such that*

$$0 < |z - p| < \varepsilon.$$

Note that p need not be in E.

Definition 1.5. *A point which is not a limit point of E is called an **isolated point** of E.*

The following gives a useful equivalent definition of limit point in terms of limits of sequences.

Theorem 1.1. *A point $p \in \mathbb{R}$ is a limit point of $E \subseteq \mathbb{R}$ if and only if there exists a sequence (p_n) in E s.t. $\lim_{n\to\infty} p_n = p$ and $p_n \neq p$, $\forall n \in \mathbb{N}$.*

Proof: See problem sheet 1.

There are all sorts of exotic examples of limit points but most sets we will consider are intervals so the following result is crucial:

Theorem 1.2. *$p \in \mathbb{R}$ is a limit point of an interval (a, b) (or $(a, b]$, or $[a, b)$ or $[a, b]$) if and only if $p \in [a, b]$.*

Proof. There are (by trichotomy) only three cases: $p < a$, $p \in [a, b]$, and $p > b$. In the first take $\varepsilon := (a - p)/2$ and get a contradiction, in the third take $\varepsilon := (p - b)/2$. If $p \in [a, b)$, given $\varepsilon > 0$ choose $x = p + \frac{1}{2}\min\{\varepsilon, (b - p)\}$. The case $p = b$ is similar. $\square$

1.3 Functions

Let $f : X \to Y$, where X and Y are subsets of $\mathbb{C}$ or $\mathbb{R}$.

Although there's no such thing a typical function here are three examples, which are often useful as test cases when we formulate definitions and make conjectures.

Example 1.1. *$f(x) = \sqrt{1-x^2}$ with domain $E = [-1, 1]$. What is its graph? Its graph looks continuous ....*

Example 1.2. *Consider function f on $E = (0, 1]$ given by*

$$f(x) := \begin{cases} \frac{p}{q} & \text{if } x = \frac{p}{q} \text{ in lowest terms,} \\ 0 & \text{when } x \text{ is irrational.} \end{cases}$$

This time our sketch of the graph is a bit more sketchy. [Try with MuPAD].

Remember that any nonempty interval contains both rational and irrational numbers. Intuitively $f(x) \to 0$ as $x \to 0$ but the limit does not exist at any other point.

Example 1.3. *The function $f(x) = x \sin \frac{1}{x}$ with domain $\mathbb{R}\backslash\{0\}$ is an important test case. As x gets close to 0, the values of f oscillate, but they do get close to 0. We will see that f has limit 0 as x goes to 0.*

1.4 Limits of Functions

Having looked at these examples we make a definition.

Definition 1.6. *Let $E \subseteq \mathbb{R}$ (or $\mathbb{C}$), and $f : E \to \mathbb{R}$ (or $\mathbb{C}$) be a real (or complex) function. Let p be a* limit point *of E and let l be a number. We say that f **tends to** l **as** x **tends to** p if $\forall \varepsilon > 0\ \exists \delta > 0$ such that $\forall x \in E$ such that $0 < |x - p| < \delta$*

$$|f(x) - l| < \varepsilon \quad .$$

In symbols we write this as '$\lim_{x\to p} f(x) = l$' or '$f(x) \to l$ as $x \to p$.'

Remark 1.3. *(i) Note that p is not necessarily in E.*

(ii) Note that in the definition δ may depend on p and ε.

Example 1.4. *Let $\alpha > 0$. Consider the function $f(x) = |x|^\alpha \sin \frac{1}{x}$ on the domain $E = \mathbb{R}\backslash\{0\}$. Show that $f(x) \to 0$ as $x \to 0$.*

Since $|\sin\theta| \leqslant 1$ we have that $\left|x^\alpha \sin \frac{1}{x}\right| \le |x|^\alpha$ for any $x \neq 0$. Therefore, $\forall\ \varepsilon > 0$, choose $\delta = \varepsilon^{1/\alpha}$. Then

$$\left|x^\alpha \sin \frac{1}{x} - 0\right| \le |x|^\alpha < \varepsilon \qquad \text{whenever } 0 < |x - 0| < \delta.$$

According to the definition, $|x|^\alpha \sin \frac{1}{x} \to 0$ as $x \to 0$.

Example 1.5. *Consider the function $f(x) = x^2$ on the domain $E = \mathbb{R}$. Let $a \in \mathbb{R}$. Show that $f(x) \to a^2$ as $x \to a$.*

Note that $|x^2 - a^2| = |x-a||x+a| \leq |x-a|(|x| + |a|)$. So we want to get a bound on x. Suppose that $|x-a| < 1$, then

$$|x| = |x - a + a| \leq |x-a| + |a| < 1 + |a|.$$

So $\forall \epsilon > 0$, choose $\delta = \min\{1, \frac{\epsilon}{1+2|a|}\}$. Then

$$|x^2 - a^2| \leq |x-a|((1+|a|) + |a|) < \epsilon \text{ whenever } |x-a| < \delta$$

as required.

Theorem 1.4. *Let $f : E \to \mathbb{R}$ (or $\mathbb{C}$) and p be a limit point of E. If f has a limit as $x \to p$, then the limit is unique.*

Proof. Suppose $f(x) \to l_1$ and also $f(x) \to l_2$ as $x \to p$, where $l_1 \neq l_2$. Then $\frac{1}{2}|l_1 - l_2| > 0$, so by definition, $\exists \delta_1 > 0$ such that $\forall x \in E$ such that $0 < |x-p| < \delta_1$

$$|f(x) - l_1| < \tfrac{1}{2}|l_1 - l_2|$$

Similarly, $\exists \delta_2 > 0$ such that $\forall x \in E$ such that $0 < |x-p| < \delta_2$

$$|f(x) - l_2| < \tfrac{1}{2}|l_1 - l_2|.$$

Let $\delta = \min\{\delta_1, \delta_2\}$. Since p is a limit point of E and $\delta > 0$, $\exists x_0 \in E$ such that $0 < |x_0 - p| < \delta$. However

$$\begin{aligned} |l_1 - l_2| &= |(f(x_0) - l_1) - (f(x_0) - l_2)| && \text{[Add and subtract technique]} \\ &\leqslant |f(x_0) - l_1| + |f(x_0) - l_2| && \text{[Triangle Law]} \\ &< \tfrac{1}{2}|l_1 - l_2| + \tfrac{1}{2}|l_1 - l_2| \\ &= |l_1 - l_2| \end{aligned}$$

a contradiction. □

Remark 1.5. *An exercise in contrapositives:* f doesn't converge *to l as $x \to p$ (i.e. either f has no limit or $f(x) \to a \neq l$ as $x \to p$), means that*
$\exists \varepsilon > 0$, such that $\forall \delta > 0$, $\exists x \in E$ such that $0 < |x-p| < \delta$ but $|f(x) - l| \geqslant \varepsilon$.

The following theorem translates questions about function limits to questions about sequence limits, and so we can make use of results in Analysis I. For example we will be able to deduce an algebra of limits for functions from the algebra of limits for sequences.

Theorem 1.6. *Let $f : E \to \mathbb{R}$ (or $\mathbb{C}$) where $E \subseteq \mathbb{R}$ (or $\mathbb{C}$), p be a limit point of E and $l \in \mathbb{C}$. Then the following two statements are equivalent:*

(a) *$f(x) \to l$ as $x \to p$;*

(b) *For every sequence (p_n) in E such that $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$ we have that $f(p_n) \to l$ as $n \to \infty$.*

Informally $f(x) \to l$ as $x \to p$ if and only if f tends to the same limit l along any sequence in E going to p.

Proof. $\Longrightarrow$: Suppose $\lim_{x\to p} f(x) = l$. Then $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $\forall x \in E$ such that $0 < |x - p| < \delta$

$$|f(x) - l| < \varepsilon.$$

Now suppose (p_n) is a sequence in E, with $p_n \to p$ and $p_n \neq p$. Then $\exists N \in \mathbb{N}$ such that $\forall n > N$

$$|p_n - p| < \delta.$$

So, since $p_n \neq p$, $\forall n > N$

$$|f(p_n) - l| < \varepsilon.$$

Hence, $\lim_{n\to\infty} f(p_n) = l$.

$\Longleftarrow$: Argue by contradiction. Suppose $\lim_{x\to p} f(x) = l$ is not true. Then $\exists \varepsilon_0 > 0$, such that $\forall \delta > 0$,—which we choose to be $1/n$ for arbitrary n— $\exists x_n \in E$, with $0 < |x_n - p| < 1/n$ but

$$|f(x_n) - l| \geqslant \varepsilon_0.$$

Therefore we have found a sequence (x_n) which converges to p but $(f(x_n))$ does not tend to l. Contradiction. □

The above result is very useful when we want to prove that limits do not exist.

Example 1.6. *Show that $\lim_{x\to 0} \sin \frac{1}{x}$ doesn't exist.*

Let $x_n = \dfrac{1}{\pi n}$ and $y_n = \dfrac{1}{2\pi n + \frac{\pi}{2}}$. Then both sequences x_n and y_n tend to 0, but

$$\lim_{n\to\infty} \sin \frac{1}{x_n} = 0$$

and

$$\lim_{n\to\infty} \sin \frac{1}{y_n} = 1.$$

So $\lim_{x\to 0} \sin \frac{1}{x}$ cannot exist.

1.5 Algebra of Limits

We can use the theorem of the previous subsection together with the Algebra of Limits of Sequences to prove the corresponding results: we get the Algebra of Limits of Functions. We state the theorem for $\mathbb{C}$ but it also holds for $\mathbb{R}$.

Theorem 1.7. *Let $E \subseteq \mathbb{C}$ and let p be a limit point of E. Let $f, g : E \to \mathbb{C}$, and let $\alpha, \beta \in \mathbb{C}$. Suppose that $f(x) \to A$, $g(x) \to B$ as $x \to p$. Then the following limits exist and have the values stated:*

(Linear Combination) $\lim_{x \to p} (\alpha \cdot f + \beta \cdot g)(x) = \alpha A + \beta B$;

(Product) $\lim_{x \to p} (f(x)g(x)) = AB$;

(Quotient) *if $B \neq 0$ then $\exists \delta > 0$ s.t. $g(x) \neq 0$ $\forall x \in E$ such that $0 < |x - p| < \delta$, and $\lim_{x \to p} (f(x)/g(x)) = A/B$;*

(Weak Inequality) *if $f(x) \geqslant 0$ for all $x \in E$ then $A \geqslant 0$.*

Proof. These can all be deduced directly from the Algebra of Limits of Sequences using Theorem 1.6, or they can be proved directly from the definitions (just mimic the sequence proofs).

Some examples:

Direct proof of product result: Note

$$|f(x)g(x) - AB| \leq |f(x)||(g(x) - B)| + |B||(f(x) - A)|,$$

so we need to bound $|f(x)|$. But $\exists \delta_1 > 0$ s.t. $\forall x \in E$ such that $0 < |x - p| < \delta_1$, $|f(x) - A| < 1$. So

$$|f(x)| \leq |f(x) - A| + |A| < 1 + |A|.$$

Now given $\varepsilon > 0$, $\exists \delta_2 > 0$ such that $\forall x \in E$ such that $0 < |x - p| < \delta_2$

$$|f(x) - A| < \frac{\varepsilon}{1 + |A| + |B|},$$

and $\exists \delta_3 > 0$ such that $\forall x \in E$ such that $0 < |x - p| < \delta_3$

$$|g(x) - B| < \frac{\varepsilon}{1 + |A| + |B|}.$$

Thus, taking $\delta = \min\{\delta_1, \delta_2, \delta_3\}$, $\forall x \in E$ such that $0 < |x - p| < \delta$

$$|f(x)g(x) - AB| \leq (1 + |A|)|g(x) - B| + |B||f(x) - A| < \varepsilon.$$

To prove: If $B \neq 0$ then $\exists \delta > 0$ s.t. $g(x) \neq 0$ $\forall x \in E$ such that $0 < |x - p| < \delta$, and $\lim_{x \to p} (1/g(x)) = 1/B$;

I will do it both ways:

(i) Deduction from AOL for sequences: Suppose first that there is no such δ. Then for each n, $\exists p_n \in E$ such that $0 < |p_n - p| < 1/n$, and $g(p_n) = 0$. But then $p_n \to p$, so $g(p_n) \to B$, giving $B = 0$, a contradiction. So $\delta > 0$ exists.

Now let (x_n) be any sequence in E with $x_n \to p$ and $x_n \neq p$. We may assume $x_n \in (p - \delta, p + \delta)$ (by tails). Hence $g(x_n) \neq 0$ and $g(x_n) \to B$. Thus by the AOL for sequences, $1/g(x_n) \to 1/B$. Thus, by Theorem 1.6, $1/g(x) \to 1/B$ as required.

(ii) Direct proof: Take $\varepsilon = |B|/2 > 0$. So $\exists \delta_1 > 0$ such that $\forall x \in E$ such that $0 < |x - p| < \delta_1$

$$|g(x) - B| < |B|/2.$$

Thus by the Triangle Law $\forall x \in E$ such that $0 < |x - p| < \delta_1$

$$|g(x)| = |B + (g(x) - B)| \geq |B| - |g(x) - B| > |B| - |B|/2 = |B|/2.$$

(So in particular, $g(x) \neq 0$ whenever $0 < |x - p| < \delta_1$.)

Now, given $\varepsilon > 0$, $\exists \delta_2 > 0$ such that $\forall x \in E$ such that $0 < |x - p| < \delta_2$

$$|g(x) - B| < |B|^2 \varepsilon/2.$$

Take $\delta = \min\{\delta_1, \delta_2\}$. Then if $x \in E$ is such that $0 < |x - p| < \delta$

$$|1/g(x) - 1/B| = \frac{|g(x) - B|}{|g(x)||B|} < \frac{|B|^2 \varepsilon/2}{|B|^2/2} = \varepsilon$$

as required. □

Corollary 1.8. *Note we have also proved above that if* $\lim_{x \to p} g(x) = B \neq 0$, *then there is a positive number* $\delta > 0$ *such that*

$$|g(x)| \geqslant \frac{|B|}{2} \qquad \forall x \in E \textit{ such that } 0 < |x - p| < \delta.$$

In particular, $|g(x)| > 0$ $\forall x \in E$ *such that* $0 < |x - p| < \delta$.

It can be proved similarly that if $g : E \to \mathbb{R}$ *and* $B > 0$, *then* $\exists \delta > 0$ *such that* $g(x) > \frac{B}{2} > 0$ $\forall x \in E$ *such that* $0 < |x - p| < \delta$.

The following is a version of the sandwich theorem.

Proposition 1.9. *Let* $E \subseteq \mathbb{R}$ *and let* p *be a limit point of* E. *Let* $f, m, M : E \to \mathbb{R}$. *Suppose that there exists* $\delta > 0$ *s.t.* $m(x) \leq f(x) \leq M(x)$ *for all* $x \in E$ *such that* $0 < |x - p| < \delta$ *and that* $m(x) \to l$, $M(x) \to l$ *as* $x \to p$. *Then* $\lim_{x \to p} f(x)$ *exists and equals* l.

For proof see problem sheet 1.

1.6 An extension

Sometimes we want to extend the notion '$f(x) \to \ell$ as $x \to p$' to cover 'infinity'. Here is one such extension: note that although ∞ appears in the language, we have **not** given it the status of a number: it can only appear in certain phrases in our mathematical language which are shorthand for quite complicated statements about real numbers.

Definition 1.7. *Suppose that $E \subseteq \mathbb{R}$ is a set which is unbounded above and $f : E \to \mathbb{R}$. Then we write $f(x) \to \infty$ as $x \to \infty$ to mean that:*
$\forall B > 0$, $\exists D > 0$ such that $\forall x \in E$ such that $x > D$

$$f(x) > B.$$

2 Continuity of functions

We all have a good informal idea of what it means to say that a function has a continuous graph: we can draw it without lifting the pencil from the paper. But we want now to use our precise definition of '$f(x) \to l$ as $x \to p$' to discuss the idea of continuity. That is we want to discuss the precise question of whether f is continuous at a particular point p.

2.1 Definition

In the definition of $\lim_{x \to p} f(x)$, the point p need not belong to the domain E of f. But even if it does, and $f(p)$ is well-defined, the limit of f at p may not be $f(p)$.

The classic example is the function

$$f(x) := \begin{cases} 0 & \text{if } x \neq 0, \\ 1 & \text{otherwise.} \end{cases}$$

Then $\lim_{x \to 0} f(x) = 0 \neq 1$.

This example motivates our definition.

Definition 2.1. *Let $f : E \to \mathbb{R}$ (or $\mathbb{C}$), where $E \subseteq \mathbb{R}$ (or $\mathbb{C}$), and $p \in E$. We say that f **is continuous at** p if $\forall \varepsilon > 0$ $\exists \delta > 0$ such that $\forall x \in E$ such that $|x - p| < \delta$*

$$|f(x) - f(p)| < \varepsilon.$$

We continue with the notation of the definition for a moment and see what this means for isolated and limit points.

Proposition 2.1. *f is continuous at any isolated point of E.*

Proof. As p is isolated there exists $\delta > 0$ such that there are no other points of $x \in E$ such that $0 < |x - p| < \delta$. The inequality required is therefore vacuously true. $\square$

Proposition 2.2. *If $p \in E$ is a limit point of E, then f is continuous at p, if and only if*

$$\lim_{x \to p} f(x) \quad \textit{exists and} \quad \lim_{x \to p} f(x) = f(p).$$

Proof. It's clear that the continuity definition implies the limit one at once. The limit one, provided the limit is $f(p)$, delivers all that we need for continuity except that the inequality $|f(x) - l| < \varepsilon$ holds for $x = p$ as well as the other points x in $|x - p| < \delta$. But this is immediate. $\square$

The following theorem follows immediately from Proposition 2.2 and the proof of Theorem 1.6. In this case we do not need to avoid sequences which hit the point:

Theorem. *1.6 ′. Let $f : E \to \mathbb{R}$ (or $\mathbb{C}$) where $E \subseteq \mathbb{R}$ (or $\mathbb{C}$) and $p \in E$. Then the following two statements are equivalent:*

(a) *$f(x)$ is continuous at p;*

(b) *For every sequence (p_n) in E such that $\lim_{n\to\infty} p_n = p$ we have that $f(p_n) \to f(p)$ as $n \to \infty$.*

2.2 Examples

Example 2.1. *Let $\alpha > 0$. The function $f(x) = |x|^\alpha \sin \frac{1}{x}$ not defined at $x = 0$ so it makes no sense to ask if it is continuous there. In such circumstances we modify f in some suitable way. So we look at*

$$g(x) := \begin{cases} |x|^\alpha \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then 0 is a limit point of the domain, and we calculated before that $\lim_{x\to 0} g(x) = 0 = g(0)$, so g is continuous at 0.

Example 2.2. *Let $f : (0, 1] \to \mathbb{R}$ be defined by*

$$f(x) := \begin{cases} \frac{1}{n} & \text{if } x = \frac{m}{n} \text{ in lowest terms,} \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

At which points of $(0, 1]$ is f continuous?

This is very like problem on the Exercise Sheets, so I won't give a full proof here, only indicate how I would tackle it.

Every $p \in (0,1]$ is a limit point, so we need to work out $\lim_{x\to p} f(x)$ for each p. We know that we can do this by looking at $\lim_{n\to\infty} f(p_n)$ for each sequence (p_n) converging to p.

We know that there is always a sequence of irrationals (x_n) converging to p. (Because, from Analysis I, for every $n \in \mathbb{N}$ the interval $(p, p+1/n)$ contains an irrational number x_n.) Then the sequence $(f(x_n))$ is just the null sequence $(0, 0, \dots)$ with limit 0.

So it looks as if we should distinguish between rational and irrational points.

Suppose $p \neq 0$ is rational. Then, with (x_n) as above, $f(x_n) \to 0$ but $f(p) \neq 0$. Therefore f is not continuous at non-zero rational points, by Theorem 1.6.

Now let p be irrational. Some sequences (for example irrational ones) tend to $0 = f(p)$. But do all sequences have this property? Let (p_n) be any sequence in $(0,1]$ tending to p and consider $f(p_n)$. If this does not tend to zero, then for some $\varepsilon > 0$ we can find a subsequence such that $f(p_{n_j}) > \varepsilon$. That is, these p_{n_j} must be rational and if $p_{n_j} = \frac{r_{n_j}}{q_{n_j}}$ in lowest terms then, as $f(p_{n_j}) = \frac{1}{q_{n_j}}$, the denominator $q_{n_j} < \frac{1}{\varepsilon}$. There are only a finite number of such points in the interval, so there exists $\delta > 0$ such that there are no such points in the interval $(p-\delta, p+\delta)$. Thus —since $p_{n_j} \to p$— we cannot have the claimed subsequence.

Therefore f is continuous at irrational points since for all sequences (p_n) we have that $f(p_n) \to f(p)$. $\square$

2.3 Algebraic properties

We can use our characterisation of continuity at limit points in terms of $\lim_{x\to p} f(x)$ together with the Algebra of Function Limits to prove that the class of functions continuous at p is closed under all the usual operations. We state the theorem for $\mathbb{C}$ but it also holds for $\mathbb{R}$.

Theorem 2.3. *Let $E \subseteq \mathbb{C}$ and let $p \in E$. Let $f, g : E \to \mathbb{C}$, and let $\alpha, \beta \in \mathbb{C}$. Suppose that f, g are continuous at p. Then the following functions are also continuous at p:*

(Linear Combination) $(\alpha \cdot f + \beta \cdot g)(x)$;

(Product) $(f(x)g(x))$; *and*

(Quotient) $(f(x)/g(x))$ *provided* $g(p) \neq 0$ *(which guarantees that there exists δ such that $f(x)/g(x)$ is defined $\forall x \in E$ such that $|x-p| < \delta$).*

Proof. Follows directly from the Algebra of Function Limits. However, it is a good[2] exercise to write out a proof from the definition—again just mimic what was done for the AOL for sequences. $\square$

Example 2.3. *Let $f : \mathbb{C} \to \mathbb{C}$ (or $\mathbb{R} \to \mathbb{R}$) be a polynomial. Then f is continuous at every point of $\mathbb{C}$ (or $\mathbb{R}$).*

Further, if $f(x) = \frac{r(x)}{q(x)}$, where $r, q : \mathbb{C} \to \mathbb{C}$ (or $\mathbb{R} \to \mathbb{R}$) are polynomials. Then if $q(p) \neq 0$, f is continuous at p.

This follows immediately from the above theorem because the function $f(x) = x$ with domain $\mathbb{C}$ (or $\mathbb{R}$) is continuous.

2.4 Composition of continuous functions

However we can do more than these trivial algebraic results.

Theorem 2.4. *Let $f : E \to \mathbb{C}$ and $g : f(E) \to \mathbb{C}$, and define $h : E \to \mathbb{C}$ by*

$$h(x) = (g \circ f)(x) := g(f(x)) \qquad \text{for } x \in E.$$

If f is continuous at $p \in E$ and g is continuous at $f(p)$, then h is continuous at p.

Proof. For any $\varepsilon > 0$, since g is continuous at $f(p)$, $\exists \delta_1 > 0$ such that $\forall y \in f(E)$ such that $|y - f(p)| < \delta_1$

$$|g(y) - g(f(p))| < \varepsilon.$$

That is $\forall x \in E$ such that $|f(x) - f(p)| < \delta_1$

$$|g(f(x)) - g(f(p))| < \varepsilon.$$

However, f is continuous at p, so $\exists \delta > 0$ such that $\forall x \in E$ such that $|x - p| < \delta$

$$|f(x) - f(p)| < \delta_1 \qquad .$$

Hence $\forall x \in E$ such that $|x - p| < \delta$

$$|g(f(x)) - g(f(p))| < \varepsilon$$

so that h is continuous at p. $\square$

[2] Doing this will reinforce the definitions, but also consolidate your understanding of sequences.

2.5 More examples of continuous functions.

Recall from Analysis I that the functions from $\mathbb{C} \to \mathbb{C}$ (or $\mathbb{R} \to \mathbb{R}$), $\exp(x)$, $\sin(x)$, $\cos(x)$, $\sinh(x)$ and $\cosh(x)$ etc are defined by their power series, each of which has infinite radius of convergence. Later we will see that a power series is continuous within its radius of convergence, so each of these functions is continuous everywhere and, for now, we will assume this. We can now use the algebra of continuous functions and the composition of continuous functions to prove the continuity of a wide variety of functions.

Example 2.4. *The function* $g : \mathbb{R} \to \mathbb{R}$

$$g(x) := \begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

is continuous at every point of $\mathbb{R}$.

Proof *We have already proved that* g *is continuous at* 0.

If $p \neq 0$*:* $1/x$ *is continuous at* p *as* $p \neq 0$ *[Quotient of continuous functions] and* $\sin(x)$ *is continuous at* $1/p$*. Hence* $\sin(1/x)$ *is continuous at* p *[Theorem 2.4]. Hence* $x \sin(1/x)$ *is continuous at* p *[Product of continuous functions].*

3 Continuity and Uniform Continuity

3.1 Continuous functions on sets

Having made our definition of 'continuity' we will see that actually, what usually matters is not continuity at a point, but continuity at all points of a set, and the interesting sets are usually intervals or disks. In the later lectures we are going to establish several important theorems about continuous functions on bounded intervals.

But here is the definition of continuity on a set.

Definition 3.1. *Let* $f : E \to \mathbb{R}$ *(or* $\mathbb{C}$*). We say that* f *is* ***continuous on*** E *if* f *is continuous at every point of* E*.*

For later use we decode this in terms of εs and δs.

Proposition 3.1. *Let* $f : E \to \mathbb{R}$ *(or* $\mathbb{C}$*). Then* f *is continuous on* E *if,*

$$\forall p \in E \text{ and } \forall \varepsilon > 0, \exists \delta > 0 \text{ such that } \forall x \in E \text{ such that } |x - p| < \delta$$

$$|f(x) - f(p)| < \varepsilon.$$

Note that the δ may depend on ε and on the point p.

We are about to look at uniform continuity, in which δ does not depend on p. First we will consider an example which is not uniformly continuous.

3.2 An Example

We look at an example of a function continuous on a set.

Example 3.1. *Let $f : (0, \infty) \to \mathbb{R}$ be given by $f(x) := \dfrac{1}{x}$.*

Show that for every $p \neq 0$, $\lim_{x \to p} \frac{1}{x} = \frac{1}{p}$, and thus $f(x)$ is continuous on $(0, \infty)$.

By the algebra of limits this is all clear. But we want to analyse what is going on more carefully, to see how the δ is related to ε and the the point x in question.

First,

$$|f(x) - f(p)| = \left| \frac{1}{x} - \frac{1}{p} \right| = \frac{|x - p|}{|x||p|}$$

and we can see that the problem term is $\dfrac{1}{x}$.

However, $|p| > 0$, and so when $|x - p| < \frac{1}{2}|p|$ we have by the Triangle Law that

$$|x| \geqslant |p| - |x - p| > \tfrac{1}{2}|p|;$$

so we're going to have to pick $\delta \leqslant \frac{1}{2}|p|$.

For these x, then, we have that

$$|f(x) - f(p)| \leqslant \frac{2}{|p|^2}|x - p|$$

and if we make sure $\frac{2}{|p|^2}|x - p| < \varepsilon$ we will be done.

This can be achieved that by choosing

$$\delta := \min\left(\tfrac{1}{2}|p|, \tfrac{1}{2}\varepsilon|p|^2\right)$$

which is indeed positive.

Note that for small ε (the interesting ones) the values of δ we need depend heavily on p. Near 1 choosing $\frac{1}{2}\varepsilon$ will do, but at 10^{-6} we need $\frac{1}{2.10^{12}}\varepsilon$. Our function is certainly continuous at every point, but there's no way of controlling over the whole interval how far it strays in a small neighbourhood.

3.3 Uniform Continuity

Sometimes we want to be able to control what happens over a set more 'uniformly'.

Definition 3.2. *Let $f : E \to \mathbb{R}$ (or $\mathbb{C}$). Then f is* ***uniformly continuous on*** *E if,*

$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ such that } \forall p \in E \text{ and } \forall x \in E \text{ such that } |p - x| < \delta$$

$$|f(p) - f(x)| < \varepsilon \qquad .$$

Note the difference[3]between this and the definition of 'continuous on E'.
In this, the uniform case, δ can depend only on ε and must be independent of x and p. (We must be able to choose it 'uniformly'.) Obviously if we can do this it is very nice, it gives us a way of controlling what happens on a set all at once.

Of course $f : E \to \mathbb{R}$ (or $\mathbb{C}$) is uniformly continuous on E implies that f is continuous on E.

Here is one class of functions that satisfy the uniform continuity condition.

Example 3.2. *Suppose that f is Lipschitz continuous in E: that is, assume that there exists $M > 0$ such that*

$$|f(x) - f(y)| \leqslant M|x - y| \qquad \forall x, y \in E.$$

Then f is uniformly continuous on E.

Take $x, y \in E$. Given $\varepsilon > 0$, choose $\delta = \frac{\varepsilon}{M+1} > 0$. Then

$$\begin{aligned} |f(x) - f(y)| &\leqslant M|x - y| \\ &\leqslant M\left(\frac{\varepsilon}{M+1}\right) < \varepsilon \end{aligned}$$

whenever $|y - x| < \delta$.

Note that our choice of δ **does not** depend on x or y. For a given $\varepsilon > 0$ we can find a δ that works for all x and y.

Example 3.3. $f(x) = \sqrt{x}$ *is Lipschitz continuous on $[1, \infty)$, so it is uniformly continuous.*

To see the Lipschitz condition note that

$$|\sqrt{x} - \sqrt{y}| \leqslant \frac{|x - y|}{\sqrt{x} + \sqrt{y}} \leqslant \tfrac{1}{2}|x - y|$$

for all $x, y \geqslant 1$.

3.4 Continuity implies Uniform Continuity on $[a, b]$

Our first real theorem is:

[3]For those who like pure formulae,

Continuity on E:	$\forall p \in E\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in E\ [|x - p| < \delta \Longrightarrow |f(x) - f(p)| < \varepsilon]$
Uniform Continuity on E:	$\forall \varepsilon > 0\ \exists \delta > 0\ \forall p \in E\ \forall x \in E\ [|x - p| < \delta \Longrightarrow |f(x) - f(p)| < \varepsilon]$

Swapping $\forall$s doesn't give problems, but swapping the $\forall p$ and $\exists \delta$ is the crunch.

Theorem 3.2 (**Uniform Continuity on** $[a, b]$)**.** *If $f : [a, b] \to \mathbb{R}$ (or $\mathbb{C}$) is continuous, then f is uniformly continuous.*

More generally, a continuous function on a closed and bounded set—'compact set' as we'll say next year—is uniformly continuous.

Proof. Suppose that f were not uniformly continuous. By the contrapositive of 'uniform continuity' there would exist $\varepsilon > 0$, such that for any $\delta > 0$—which we choose as $\delta = \frac{1}{n}$ for arbitrary n—there exists a pair of points x_n, $y_n \in [a, b]$, such that

$$|x_n - y_n| < \tfrac{1}{n} \qquad \text{but} \qquad |f(x_n) - f(y_n)| \geqslant \varepsilon.$$

Since $\{x_n : n \in \mathbb{N}\} \subseteq [a, b]$ is bounded, by the Bolzano–Weierstrass Theorem there exists a subsequence (x_{n_k}) which converges to some p. Hence p must be a limit point of $[a, b]$, so $p \in [a, b]$. But

$$\begin{aligned} |y_{n_k} - p| &\leqslant |x_{n_k} - y_{n_k}| + |x_{n_k} - p| \\ &< \frac{1}{n_k} + |x_{n_k} - p| \to 0 \end{aligned}$$

Thus $x_{n_k} \to p$ and $y_{n_k} \to p$, so that by continuity at p we have

$$0 < \varepsilon \leqslant |f(x_{n_k}) - f(y_{n_k})| \leq |f(x_{n_k}) - f(p)| + |f(y_{n_k}) - f(p)| \to 0 \text{ as } k \to \infty,$$

by Theorem 1.7 $'$ (in the case continuous functions we omit the requirement that the sequence does not hit p). This gives a contradiction. □

3.5 An example on an unbounded interval

Example 3.4. *$f(x) = \sqrt{x}$ is uniformly continuous in the unbounded interval $[0, +\infty)$.*

We do this in three steps: we prove uniform continuity on $[0, 1]$, we prove uniform continuity on $[1, +\infty)$, and we patch these together.

It is easy to get that $\sqrt{x}$ is continuous on $[0, 1]$: Continuity at 0 is easy. Otherwise, provided $|x - p| < \frac{1}{2}p$ we will get

$$\left|\sqrt{x} - \sqrt{p}\right| \leqslant \frac{|x - p|}{\sqrt{x} + \sqrt{p}} \leqslant \frac{2}{3\sqrt{p}}\,|x - p|$$

and can argue from there. Thus it must be uniformly continuous by Theorem 3.2.

Secondly we have already shown that $\sqrt{x}$ is Lipschitz continuous on $[1, \infty)$, so it is uniformly continuous on $[1, \infty)$.

Now we have to patch these together. This is a standard sort of argument which we do this time as an example.

We have that for all $\varepsilon > 0$, $\exists \delta_1 > 0$ such that $\forall x, y \in [0, 1]$ such that $|x - y| < \delta_1$

$$|\sqrt{x} - \sqrt{y}| < \tfrac{1}{2}\varepsilon$$

and $\exists \delta_2 > 0$ such that $\forall x, y \in [1, \infty)$ such that $|x - y| < \delta_2$

$$|\sqrt{x} - \sqrt{y}| < \tfrac{1}{2}\varepsilon.$$

Choose $\delta = \min\{\delta_1, \delta_2\} > 0$. Then, suppose that $|x - y| < \delta$. If $x, y \geqslant 1$ or $x, y \leqslant 1$ we are done.

So suppose that $x \in [0, 1]$ and $y \geqslant 1$. Then $|x - 1| < \delta$ and $|y - 1| < \delta$ so that

$$\begin{aligned} |\sqrt{x} - \sqrt{y}| &\leqslant |\sqrt{x} - \sqrt{1}| + |\sqrt{y} - \sqrt{1}| \\ &< \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon \end{aligned}$$

Hence we have

$$|\sqrt{x} - \sqrt{y}| < \varepsilon$$

whenever x, $y \in [0, \infty)$ such that $|x - y| < \delta$. By definition, $f(x) = \sqrt{x}$ is uniformly continuous in the *unbounded* interval $[0, +\infty)$.

3.6 A counterexample on a half open interval

The condition that the interval $[a, b]$ is closed cannot be relaxed.

Example 3.5. *$f(x) = \frac{1}{x}$ is not uniformly continuous in the half open interval $(0, 1]$. (see also Example 3.2.1)*

Take $\varepsilon = 1$. We show that there is no $\delta > 0$ such that definition 3.3.1 holds.

Take sequences $x_n = \frac{1}{n}$ and $y_n = \frac{1}{n+1}$. Then $|f(x_n) - f(y_n)| = 1$, but $|x_n - y_n| \to 0$. So for any $\delta > 0$, there exists n such that $|x_n - y_n| < \delta$ but $|f(x_n) - f(y_n)| \not< 1$. So f is not uniformly continuous.

4 Continuous functions on a closed and bounded interval

4.1 Boundedness

We begin with some definitions.

Definition 4.1. *Let $f : E \to \mathbb{R}$ (or $\mathbb{C}$). We say that f is **bounded on** E if $\exists M \geq 0$ such that $\forall z \in E$*

$$|f(z)| \leqslant M \qquad .$$

*We also say that f is **bounded by** M **on** E and M is a **bound** for f on E.*

Here is one of the central theorems of the course:

Theorem 4.1 (Continuous functions on $[a,b]$ are bounded). *If $f : [a,b] \to \mathbb{R}$ (or $\mathbb{C}$) is continuous, then f is bounded.*

Proof. Argue by contradiction. Suppose f were unbounded, then for any $n \in \mathbb{N}$, there exists $x_n \in [a,b]$ such that $|f(x_n)| \geqslant n$. Since (x_n) is bounded, by the Bolzano–Weierstrass Theorem, there exists a subsequence (x_{n_k}) converging to p, say. Then p is a limit point of the interval $[a,b]$ so $p \in [a,b]$. Now f is continuous at p and so we have that

$$f(p) = \lim_{k \to \infty} f(x_{n_k})$$

so in particular the sequence $(f(x_{n_k}))$ is convergent. Hence, by an Analysis I result, this sequence is bounded. But $|f(x_{n_k})| \geqslant n_k \geqslant k$. so this is a contradiction.

Therefore f must be bounded. □

Example: The function $f(x) = \frac{1}{x}$ is continuous but not bounded on $(0,1]$, so the condition that the interval is closed is required.

We will now show that these bounds are 'attained'.

Notation 4.2. *Let $f : E \to \mathbb{R}$ be a bounded real-valued function, with $E \neq \emptyset$. Then write*

$$\sup_{x \in E} f(x) := \sup\{f(t) \mid t \in E\}$$
$$\inf_{x \in E} f(x) := \inf\{f(t) \mid t \in E\}$$

noting that these exist by the Completeness Axiom.

Corollary 4.2. *Let $f : [a,b] \to \mathbb{R}$ be continuous then $\sup_{x \in [a,b]} f(x)$ and $\inf_{x \in [a,b]} f(x)$ exist.*

Proof. Immediate. □

Note 4.3. *Recall that the supremum is precisely this: an upper bound, such that nothing smaller is an upper bound. It is convenient to translate this into ε-language about functions as follows:*

$$M = \sup_{x \in E} f(x) \quad \textit{if and only if} \quad \begin{cases} \forall x \in E,\ f(x) \leqslant M;\ \textit{and} \\ \forall \varepsilon > 0\ \exists x_\varepsilon \in E \ \textit{such that}\ f(x_\varepsilon) > M - \varepsilon. \end{cases}$$

We have a similar characterisation of infimum:

$$m = \inf_{x \in E} f(x) \quad \textit{if and only if} \quad \begin{cases} \forall x \in E,\ f(x) \geqslant m;\ \textit{and} \\ \forall \varepsilon > 0\ \exists x_\varepsilon \in E \ \textit{such that}\ f(x_\varepsilon) < m + \varepsilon. \end{cases}$$

Here now is our second important theorem; note that it is only for real-valued functions.

Theorem 4.4 (**Continuous functions on** $[a, b]$ **attain their bounds**). *Let* $f : [a, b] \to \mathbb{R}$ *be continuous, then* f *attains (or achieves) its supremum and infimum. That is, there exist points*[4] x_1 *and* x_2 *in* $[a, b]$ *such that* $f(x_1) = \sup_{x\in[a,b]} f(x)$ *and* $f(x_2) = \inf_{x\in[a,b]} f(x)$.

Proof. (*1st Proof: by contradiction.*) Let us prove by contradiction that the supremum M of f is attained.

Assume the contrary, that is

$$f(t) < M \quad \text{for all } t \in [a, b].$$

Consider the function g defined on $[a, b]$ by

$$g(x) = \frac{1}{M - f(x)}$$

which is positive and continuous on $[a, b]$. Therefore g is, as we have proved, bounded on $[a, b]$, by M_0 say:

$$\frac{1}{M - f(x)} = g(x) \leqslant M_0.$$

It follows that

$$f(x) \leqslant M - \frac{1}{M_0}$$

for all $x \in [a, b]$ which is a contradiction to the fact that M is the *least* upper bound.

A similar argument deals with the infimum, or apply what we have done to $-f$ and get the result at once since

$$\inf\{t \mid t \in E\} = -\sup\{-t \mid t \in E\}.$$

$\square$

As this is such an important theorem we give an alternative proof.

Proof. (*2nd Proof: sequence argument.*) The continuous function f is bounded by our earlier theorem, so that $M := \sup_{x\in[a,b]} f(x)$ exists by the Completeness Axiom of the real number system [Analysis I]. Apply the characterisation of supremum we have given, taking $\varepsilon := \frac{1}{n}$ to find a point $x_n \in [a, b]$ such that

$$M - \tfrac{1}{n} < f(x_n) \leqslant M.$$

[4] Note that x_1, x_2 may be not unique.

Since (x_n) is bounded, by the Bolzano–Weierstrass Theorem there exists a subsequence (x_{n_k}) converging to p, say. Then p is a limit point of $[a, b]$ so $p \in [a, b]$. Since f is continuous at p, we have that $f(x_{n_k}) \to f(p)$. But from the inequality

$$M - \frac{1}{n_k} < f(x_{n_k}) \leqslant M$$

we can deduce from the sandwich theorem that $f(x_{n_k}) \to M$ as $k \to \infty$. Hence, by the uniqueness of limits, $f(p) = M = \sup_{x \in [a,b]} f(x)$.

A similar argument will deal with the infimum. □

Example: Consider the function $f(x) = x^2$ for $x \in (0, 1]$. On $(0, 1]$ this is bounded and attains it supremum, but does not attain its infimum.

4.2 A Generalisation

In the proofs we have used only:
(i) $[a, b]$ is bounded;
(ii) $[a, b]$ is closed (i.e. $[a, b]$ contains all limit points of $[a, b]$;
(iii) f is continuous.

This prompts us to make the following definition:

Definition 4.3. *A subset A of $\mathbb{R}$ (or of $\mathbb{C}$) is* ***compact*** *if it is bounded, and if it contains all its limit points.*

Our proofs would then give the more general result:

Theorem. *Let $f : E \to \mathbb{R}$ be a continuous real valued function on a compact subset E of $\mathbb{R}$ or $\mathbb{C}$. Then f is bounded, uniformly continuous, and attains its bounds.*

5 The Intermediate Value Theorem

So far we have concentrated on extreme values, the supremum and the infimum. What can we say about possible values between these?

Theorem 5.1 (IVT). *Let $f : [a, b] \to \mathbb{R}$ be continuous, and let c be a number between $f(a)$ and $f(b)$. Then there is at least one $\xi \in [a, b]$ such that $f(\xi) = c$.*

This is one of the most important theorems in this course.

Proof. By considering $-f$ instead of f if necessary, we may assume that $f(a) \leqslant c \leqslant f(b)$. The cases $c = f(a)$ and $c = f(b)$ are trivial, so assume $f(a) < c < f(b)$.

Define $g(x) = f(x) - c$. Then $g(a) < 0 < g(b)$. Hence it is sufficient to prove that if $g : [a, b] \to \mathbb{R}$ is continuous and $g(a) < 0 < g(b)$, then there exists $\xi \in (a, b)$ such that $g(\xi) = 0$.

Define $E = \{x \in [a, b] : g(x) < 0\}$.

Then, $a \in E$ so $E \neq \emptyset$ and E is bounded by b. So, by the Completeness Axiom, $\xi = \sup E$ exists. Since $a \in E$ we have $\xi = \sup E \geq a$ and since b is an upper bound for E we have $\xi = \sup E \leq b$. We now prove, by contradiction, that $g(\xi) = 0$.

Suppose first that $g(\xi) < 0$ (so $\xi \in [a, b)$). Let $\varepsilon = -g(\xi) > 0$. Then $\exists \delta > 0$ such that if $x \in [a, b]$ and $|x - \xi| < \delta$ then $|g(x) - g(\xi)| < \varepsilon$ and hence $g(x) < g(\xi) + \varepsilon = 0$. Thus $[\xi, \min\{\xi + \delta, b\}) \subseteq E$, which contradicts that $\xi = \sup E$ since there exists $x \in E$ such that $x > \xi$.

Suppose now that $g(\xi) > 0$ (so $\xi \in (a, b]$). Let $\varepsilon = g(\xi) > 0$. Then $\exists \delta > 0$ such that if $x \in [a, b]$ and $|x - \xi| < \delta$ then $|g(x) - g(\xi)| < \varepsilon$ and hence $g(x) > g(\xi) - \varepsilon = 0$. Thus $(\max\{a, \xi - \delta\}, \xi] \cap E = \emptyset$, which contradicts that $\xi = \sup E$ since there is no $x \in E$ such that $x > \max\{a, \xi - \delta\}$.

Hence $g(\xi) = 0$ as required. It is now clear that $\xi \in (a, b)$. $\square$

Remark 5.2. *The proof of IVT requires more than what we needed for boundedness and the attainment of bounds. We have used the fact that $[a, b]$ is unbroken. That is, we have used the fact that $[a, b]$ is "connected".*

This proof may seem familiar from Analysis I, where a proof similar to this was used to prove the existence of $\sqrt{2}$. In fact we can now prove this directly from the IVT.

Example 5.1. *There exists a unique positive number ξ s.t. $\xi^2 = 2$.*

Proof: Consider $f(x) = x^2 - 2$. Note that $f(0) = -2$ and $f(2) = 2$. So $f : [0, 2] \to \mathbb{R}$, $f(0) < 0 < f(2)$ and also, as f is a polynomial, it is continuous. Thus, by the IVT, there exists $\xi \in (0, 2)$ such that $f(\xi) = 0$, as required. Uniqueness can be proved as in Analysis I.

More generally the IVT is often used to show that algebraic equations have solutions. In the following, if you draw the graphs of $y = e^x$ and $y = \alpha x$, you will see that if $\alpha = e$ the curves touch, if $\alpha < e$ they do not meet, but if $\alpha > e$ then they meet twice. The following example shows how to make this graphical argument rigorous using the IVT. It shows that if $\alpha > e$ there exist two solutions. Once we have covered differentiability you will be able to prove that there are exactly two solutions, by using the fact that $f'(x) < 0$ if $x < \log\alpha$, but $f'(x) > 0$ if $x > \log\alpha$.

Example 5.2. *Let $\alpha > e$. Show that there exist two distinct points $x_i > 0$, $i = 1, 2$, such that $e^{x_i} = \alpha x_i$.*

Proof: Consider $f(x) = e^x - \alpha x$. We will prove later that for all x, e^x is continuous. Hence $f(x)$ is continuous on $[0, \infty)$. e^x is defined by its power series so that $e^x > \frac{x^2}{2}$. Thus $e^X > \alpha X$ for any $X > 2\alpha$. Fix such an $X (> \log\alpha)$.

Then $f(0) = 1 > 0$, $f(\log\alpha) = \alpha(1 - \log\alpha) < 0$, $f(X) > 0$. So we can apply the IVT to the two intervals $[0, \log\alpha]$, and $[\log\alpha, X]$ to find that there exist $x_1 \in [0, \log\alpha]$ such that $f(x_1) = 0$, and $x_2 \in [\log\alpha, X]$ such that $f(x_2) = 0$ as required.

Here (for interest) is a sketch of an alternative proof, which identifies ξ by repeated bisection.

Alternative proof to IVT. By considering $-f$ instead of f if necessary, we may assume that $f(a) \leqslant c \leqslant f(b)$. The cases $c = f(a)$ and $c = f(b)$ are trivial, so assume $f(a) < c < f(b)$.

Define $g(x) = f(x) - c$. Then $g(a) < 0 < g(b)$.

Let $x_1 = a$ and $y_1 = b$. Divide the interval $[x_1, y_1]$ into two equal parts.

If $g(\frac{1}{2}(x_1 + y_1)) = 0$ then $\xi := \frac{1}{2}(x_1 + y_1)$ will do.

Otherwise, if $g(\frac{1}{2}(x_1 + y_1)) > 0$, we choose $x_2 = x_1$ and $y_2 = (\frac{1}{2}(x_1 + y_1)$,

or, if $g(\frac{1}{2}(x_1 + y_1)) < 0$, we choose $x_2 = \frac{1}{2}(x_1 + y_1)$ and $y_2 = y_1$.

Then

$$g(x_2)g(y_2) < 0; \quad [x_2, y_2] \subset [x_1, y_1]; \text{ and } |y_2 - x_2| = \tfrac{1}{2}(y_1 - x_1).$$

Apply the same argument to $[x_2, y_2]$ instead of $[x_1, y_1]$, we then find that: either $g(\frac{1}{2}(x_2 + y_2)) = 0$ and we can take $\xi := \frac{1}{2}(x_2 + y_2)$, or there exist x_3, y_3 such that

$$g(x_3)g(y_3) < 0; \quad [x_3, y_3] \subset [x_2, y_2]; \text{ and } |y_3 - x_3| = \tfrac{1}{2}(y_2 - x_2).$$

By repeating the same procedure, we thus find two sequences x_n, y_n, such that

(i) either $g(\frac{1}{2}(x_{n-1} + y_{n-1})) = 0$ and we can take $\xi := \frac{1}{2}(x_{n-1} + y_{n-1})$, or $g(x_n)g(y_n) < 0$;

(ii) $[x_n, y_n] \subset [x_{n-1}, y_{n-1}]$ for any $n = 2, \ldots$;

(iii) $|y_n - x_n| = \frac{1}{2}|y_{n-1} - x_{n-1}| = \cdots = \frac{1}{2^{n-1}}|y_1 - x_1| = \dfrac{b-a}{2^{n-1}}$.

Obviously, (x_n) is a bounded increasing sequence, and (y_n) is a bounded decreasing sequence. Bounded monotone sequences converge and so $x_n \to \xi$ and $y_n \to \xi'$ for some $\xi,\ \xi' \in [a,b]$. Since by Algebra of Limits

$$|\xi' - \xi| = \lim_{n\to\infty} |y_n - x_n| = \lim_{n\to\infty} \tfrac{1}{2^{n-1}}\,(b-a) = 0,$$

we get $\xi = \xi'$. Since g is continuous at ξ, we have by Algebra of Limits and the preservation of weak inequalities that

$$g(\xi)^2 = \lim_{n\to\infty} g(x_n) \lim_{n\to\infty} g(y_n) = \lim_{n\to\infty} g(x_n)g(y_n) \leqslant 0.$$

Hence $g(\xi)^2 = 0$ as we are dealing with real numbers, so that $g(\xi) = 0$.

That is, $f(\xi) = c$. □

Remark 5.3. *The above proof of the IVT also provides a method of finding roots to $f(\xi) = c$, but other methods may find roots faster if additional information about f (e.g. that f is differentiable) is available.*

Corollary 5.4. *Let $\{[x_n, y_n])$ be a decreasing net*[5] *of closed intervals of $\mathbb{R}$ such that the length $y_n - x_n \to 0$. Then $\cap_{n=1}^{\infty}[x_n, y_n]$ contains exactly one point.*

Proof. Just extract the relevant lines of the IVT proof above. □

5.1 Closed bounded intervals map onto closed bounded intervals

We can reformulate the theorems of sections 4 and 5 as the following very useful theorem.

Theorem 5.5. *Let $f : [a,b] \to \mathbb{R}$ be a real valued continuous function. Then $f([a,b]) = [m, M]$ for some $m, M \in \mathbb{R}$.*

That is, a continuous real-valued function maps a closed and bounded interval onto a closed and bounded interval.

Proof. Let $m := \inf_{x\in[a,b]} f(x)$ and $M := \sup_{x\in[a,b]} f(x)$. These exist by the theorem on boundedness. Clearly $f([a,b]) \subseteq [m, M]$.

By the theorem on the attainment of bounds, there exist $\xi \in [a,b]$ and $\eta \in [a,b]$ such that $f(\xi) = m$ and $f(\eta) = M$; hence $m, M \in f([a,b])$.

Now let $y \in [m, M]$, so $f(\xi) \leq y \leq f(\eta)$. By applying the IVT to f restricted to the interval $[\xi, \eta]$ (or $[\eta, \xi]$ as case may be) we find an $x \in [\xi, \eta] \subseteq [a,b]$ such that $f(x) = y$; hence $y \in f([a,b])$. Hence $[m, M] \subseteq f([a,b])$. □

[5] That is $[x_{n+1}, y_{n+1}] \subset [x_n, y_n]$ for each n

6 Monotone Functions and the Continuous Inverse Function Theorem

6.1 Monotone Functions

The following definitions require the ordered structure of real numbers, and so apply only in $\mathbb{R}$.

Definition 6.1. *Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$. We say that:*

(a) *(i) f is **increasing** if $f(x) \leqslant f(y)$ whenever $x \leqslant y$.*

*(ii) f is **strictly increasing** if $f(x) < f(y)$ whenever $x < y$.*

(b) *(i) f is **decreasing** if $f(x) \geqslant f(y)$ whenever $x \leqslant y$.*

*(ii) f is **strictly decreasing** if $f(x) > f(y)$ whenever $x < y$.*

*A function is called **monotone** on E if it is increasing or decreasing on E.*

6.2 Continuity of the Inverse Function

Recall that the inverse function was defined in 'Introduction to Pure Mathematics' last term.

Definition 6.2. *Let $f : A \to B$ be a function. We say that 'f is invertible' if there exists a function $g : B \to A$ such that $g(f(x)) = x$ for all $x \in A$ and $f(g(y)) = y$, for all $y \in B$. We then call g an inverse of f.*

We have seen that continuous functions map intervals to intervals. We want to say something about the inverse function when it exists. Note that any result about increasing functions f can be translated into a result about decreasing functions simply by considering the functions $-f$.

We will prove:

Theorem 6.1 (**Continuous Inverse Function Theorem** (IFT))**.** *Let f be a strictly increasing and continuous real valued function on $[a, b]$. Then f has a well-defined continuous inverse on $[f(a), f(b)]$.*

This is contained in the following theorem.

Theorem 6.2. *Let $f : [a, b] \to \mathbb{R}$ be strictly increasing and continuous on $[a, b]$. Then*

(i) *$f([a, b]) = [f(a), f(b)]$;*

(ii) *f has a unique inverse $g : [f(a), f(b)] \to \mathbb{R}$;*

(iii) *g is strictly increasing;*

(iv) *g is continuous.*

Proof. (i) This is just Theorem 5.5 as in this case $m = f(a)$ and $M = f(b)$.

(ii) This is straightforward; $f : [a, b] \to [f(a), f(b)]$ is now 1—1 and onto. So given $y \in [f(a), f(b)]$ there exists a unique $x \in [a, b]$ such that $f(x) = y$. Define $g(y) = x$. So the inverse function exists and is unique.

(iii) This is also straightforward. Assume there exist $u, v \in [f(a), f(b)]$ with $u < v$ but $g(u) \geqslant g(v)$. But as f is strictly increasing this implies $u = f(g(u)) \geqslant f(g(v)) = v$, a contradiction.

(iv) We wish to prove that for any $y_0 \in [f(a), f(b)]$ the function g is continuous at y_0. For $y_0 \in (f(a), f(b))$: Given $\epsilon > 0$, if necessary take ϵ smaller such that $g(y_0)+\epsilon \in [a, b]$ and $g(y_0) - \epsilon \in [a, b]$.
Choose $\delta = \min\{f(g(y_0) + \epsilon) - y_0, y_0 - f(g(y_0) - \epsilon)\}$. (Draw the graph of $g(y)$ to see why we choose it like this) Then

$$
\begin{array}{rcl}
 & y_0 - \delta & < y < y_0 + \delta \\
\implies & f(g(y_0) - \epsilon) & < y < f(g(y_0) + \epsilon) \\
\implies & g(f(g(y_0) - \epsilon)) & < g(y) < g(f(g(y_0) + \epsilon)) \\
\implies & g(y_0) - \epsilon & < g(y) < g(y_0) + \epsilon
\end{array}
$$

and g is continuous at y_0 as required. The points $y_0 = f(a)$ and $y_0 = f(b)$ are similar.

[For example, if $y_0 = f(a)$: Given $\epsilon > 0$, if necessary take ϵ smaller such that $a+\epsilon \leq b$. Choose $\delta = f(a + \epsilon) - f(a)$. Then $y \in [f(a), f(b)]$ with $|y - f(a)| < \delta$

$$
\begin{array}{rl}
\implies & f(a) \leq y < f(a) + \delta \\
\implies & f(a) \leq y < f(a + \epsilon) \\
\implies & a \leq g(y) < a + \epsilon
\end{array}
$$

and, as $g(f(a)) = a$, g is continuous at $f(a)$ as required.] $\square$

Remark 6.3. *Note that from Q3, problem sheet 3, if $f : [a, b] \to \mathbb{R}$ is a continuous, 1–1 function with $f(a) < f(b)$, then f is strictly increasing on $[a, b]$. So for the Inverse Function Theorem (IFT) it is sufficient to assume that $f : [a, b] \to \mathbb{R}$ is continuous and 1–1.*

Note 6.4. *If you choose to use the notation f^{-1} for the inverse function then you must make very clear what you intend the domains of f and f^{-1} to be. For example sine and cosine are only invertible on a part of their domain where they are increasing or decreasing.*

6.3 Exponentials, Logarithms, Powers etc.

In the following I will consider the functions only on real domains. Some of the results extend to complex domains.

Recall from Analysis I that functions such as $\exp(x)$, $\sin(x)$, $\cos(x)$, $\sinh(x)$ and $\cosh(x)$ etc are defined by their power series each of which has infinite radius of convergence. Later we will see that a power series is continuous within its radius of convergence so each of these functions is continuous on $\mathbb{R}$. For each of them, if we take as domain a closed interval on which the function is strictly monotone, then we can use the IFT to show the function, with the given domain, has a continuous inverse. (See also Problem sheet 3 Q5)

In particular we can therefore define the exponential function: $\exp : \mathbb{R} \to \mathbb{R}$ as $\exp(x) = \sum \frac{x^n}{n!}$. Most of the following properties were proved in Analysis I (though some used results to be proved in this course):

1. $\exp'(x) = \exp(x)$;
2. $\exp(x)\exp(y) = \exp(x+y)$;
3. $\exp 0 = 1$ and $\exp(-x) = 1/\exp(x)$;
4. $\exp(x) > 0$;
5. As noted above exp is continuous. But we can also prove it directly

   Lemma 6.5. *The function* exp *is continuous.*

   Proof. We have

   $$|\exp(x+h) - \exp(x)| = \exp(x)|\exp(h) - 1)|$$

   so for $|h| < 1$ we have by the Triangle Law and the preservation of $\leqslant$ under limits

   $$|\exp(x+h) - \exp(x)| \leqslant \exp(x)\sum_{n\geqslant 1}|h|^n/n! \leqslant \exp(x)\sum_{n\geqslant 1}|h|^n = \frac{|h|}{1-|h|}\exp(x),$$

   which tends to 0 as $h \to 0$.

   $\square$

6. We can obtain numerous inequalities: For example if $x > 0$,

   $$\exp(x) = 1 + x + \frac{x^2}{2!} + \lim_{n\to\infty}\sum_{r=3}^{n}\frac{x^r}{r!} > 1 + x,$$

   and hence also if $x > 0$,

   $$\exp(-x) < \frac{1}{1+x}.$$

Remark 6.6. *Recall that in the limit strict inequalities become weak, so we have used the term* $\frac{x^2}{2!}$ *to ensure strict inequality. This useful trick was also used in Analysis I and we will use it in future without comment.*

7. The logarithm:

   Lemma 6.7. exp *is strictly increasing and* $\exp : \mathbb{R} \to (0, \infty)$ *is a bijection and hence invertible. The inverse is denoted by* $\log : (0, \infty) \to \mathbb{R}$*. Furthermore, for every* $y > 0$ *the function* log *is continuous at* y*.*

   Proof. We will prove later that exp is strictly increasing and hence 1-1 in Proposition 13.3 (using the fact that $\exp'(x) > 0\ \forall x \in \mathbb{R}$).

   Now we prove that $\exp(x)$ maps $\mathbb{R}$ *onto* $(0, \infty)$. Given $y > 0$ we can find A such that $1/(1+A) < y < 1+A$. Hence $\exp(-A) < 1/(1+A) < y < 1+A < \exp(A)$. Thus by the IVT there exists $x \in (-A, A)$ such that $\exp(x) = y$, as required.

   Finally we can apply the IFT to $\exp : [-A, A] \to [\exp(-A), \exp(A)]$. The image interval then contains y so log is continuous at y. □

   Remark 6.8. *When dealing with inverses of functions on unbounded intervals one generally proceeds in this way. First prove the function is a bijection - generally using the IVT- then show the inverse is continuous at a general point y by applying the IFT to a suitable closed bounded interval.*

8. Let e denote the real number $\mathrm{e} = \exp(1) = \sum \frac{1}{n!}$ then $\log e = 1$;

9. For any $a > 0$ and any $x \in \mathbb{R}$ we define

   $$a^x := \exp(x \log a).$$

   Then $a^{x+y} = a^x a^y$; Also $e^x = \exp(x)$;

Note: We can also define $\exp : \mathbb{C} \to \mathbb{C}$ by $\exp(z) = \sum \frac{z^n}{n!}$. The first 3 of the above properties also hold in $\mathbb{C}$ and also $\exp(z) \neq 0$.

6.4 Left-hand and Right-hand limits

For functions defined on an interval, we may talk about right-hand and left-hand limits.

Definition 6.3. *(i) Let* $f : [a, b) \to \mathbb{R}$ *(or* $\mathbb{C}$*) and* $p \in [a, b)$*; and let* $l \in \mathbb{R}$ *(or* $l \in \mathbb{C}$*). We say that* l ***is the right-hand limit of*** f ***at*** p *if,* $\forall \varepsilon > 0$*,* $\exists \delta > 0$ *such that* $\forall x \in [a, b)$ *such that* $0 < x - p < \delta$

$$|f(x) - l| < \varepsilon.$$

We write this as

$$\lim_{x \to p+} f(x) = l; \ \textit{or as} \ \lim_{\substack{x \to p \\ x > p}} f(x); \ \textit{or sometimes as} \ f(p+) = l.$$

Similarly we have:

(ii) Let $f : (a, b] \to \mathbb{R}$ (or $\mathbb{C}$) and $p \in (a, b]$; and let $l \in \mathbb{R}$ (or $l \in \mathbb{C}$). We say that l ***is the left-hand limit of f at*** *p if, $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $\forall x \in [a, b)$ such that $-\delta < x - p < 0$*

$$|f(x) - l| < \varepsilon.$$

We write this as

$$\lim_{x \to p-} f(x) = l; \ \textit{or as} \ \lim_{\substack{x \to p \\ x < p}} f(x); \ \textit{or sometimes as} \ f(p-) = l.$$

The following provides good practice in using the definitions.

Proposition 6.9. *Let $f : (a, b) \to \mathbb{C}$ and let $p \in (a, b)$. Then the following are equivalent:*

(i) $\lim_{x \to p} f(x) = l$*;*

(ii) *Both $\lim_{x \to p+} f(x) = l$ and $\lim_{x \to p-} f(x) = l$.*

Example 6.1. *Consider function $f : \mathbb{R} \to \mathbb{R}$ given by*

$$f(x) = \begin{cases} x & \textit{if } x \geqslant 0; \\ x + 1 & \textit{if } x < 0. \end{cases}$$

Then $f(0+) = 0$ and $f(0-) = 1$. But $\lim_{x \to 0} f(x)$ does not exist.

6.5 Left-continuity and Right-continuity

We translate the above definitions into 'continuity' language.

Definition 6.4. *(i) We say f is* ***right continuous at*** *p* ***if*** *$f(p+) = f(p)$.*[6]

(ii) We say f is ***left continuous at*** *p* ***if*** *$f(p-) = f(p)$.*

Again, for practice prove the following.

Proposition 6.10. *Let $f : (a, b) \to \mathbb{R}$ and let $p \in (a, b)$. Then the following are equivalent:*

(i) *f is continuous at p;*

[6]Note that we are saying that the limit exists and that it equals $f(p)$.

(ii) *f is both left-continuous at p and right-continuous at p.*

Example 6.2. *Again consider the function*

$$f(x) = \begin{cases} x & \text{if } x \geqslant 0; \\ x+1 & \text{if } x < 0. \end{cases}$$

Then at 0 f is right continuous but not left continuous. It is not continuous at 0.

6.6 Continuity of Monotone Functions

This section will be omitted from lectures and is included as an example.

We now discuss the continuity of monotone functions. Remember that any result about increasing functions f can be translated into a result about decreasing functions by considering instead the functions $-f$.

Theorem 6.11. *Let $f : (a, b) \to \mathbb{R}$ be an increasing function. Then for every $x_0 \in (a, b)$ the right-hand limit $f(x_0+)$ and the left-hand limit $f(x_0-)$ of f at x_0 exist.*

Moreover, $f(x_0-) = \sup_{a<x<x_0} f(x)$, $f(x_0+) = \inf_{x_0<x<b} f(x)$ and

$$f(x_0-) \leqslant f(x_0) \leqslant f(x_0+).$$

Proof. By hypothesis,$\{f(x) : a < x < x_0\}$ is non-empty and is bounded above by $f(x_0)$, and therefore has a least upper bound $A := \sup_{a<x<x_0} f(x)$. Then $A \leqslant f(x_0)$. We have to show that $f(x_0-) = A$. Let $\varepsilon > 0$ be given. It follows from the definition of $\sup_{a<x<x_0} f(x)$, that there is a $x_\varepsilon \in (a, x_0)$ such that

$$A - \varepsilon < f(x_\varepsilon) \leqslant A.$$

As $x_0 - x_\varepsilon > 0$ choose $\delta := x_0 - x_\varepsilon$. Then, $x \in (x_\varepsilon, x_0)$ if and only if $0 < x_0 - x < \delta$, and thus, as f is increasing

$$A - \varepsilon < \left[f(x_\varepsilon) \leqslant \right] f(x) \leqslant A \quad \text{for all } 0 < x_0 - x < \delta.$$

By definition $f(x_0-) = A$ and we are done.

The other inequality can be obtained by a similar argument (a good exercise); or by applying what we have done to the function $-f(b-x)$ on $(0, b-a)$ and juggling with the inequalities. $\square$

Remark 6.12. *Informally we call the difference $f(x_0+) - f(x_0-)$ the "jump" of f at x_0.*

7 Limits at infinity and infinite limits

7.1 Limits at infinity: functions of a real variable

We want to extend our definition of the limit '$\lim_{x\to a} f(x)$' to allow us to talk about the end points of infinite intervals like $(0, \infty)$.

Definition 7.1. *Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$ (or $\mathbb{C}$) and let $l \in \mathbb{R}$ (or $\mathbb{C}$). Suppose that for every $b \in \mathbb{R}$ the set $E \cap (b, +\infty)$ is non-empty. We say that $f(x) \to l$ **as** $x \to +\infty$ if, $\forall \varepsilon > 0$, $\exists B > 0$ such that $\forall x \in E$ such that $x > B$*

$$|f(x) - l| < \varepsilon.$$

We write this as $\lim_{x\to+\infty} f(x) = l$.

Exercise 7.1. *Make a similar definition for $\lim_{x\to-\infty} f(x) = l$.*

Note 7.1. *We will often just write '$f(x) \to l$ as $x \to \infty$' for '$f(x) \to l$ as $x \to +\infty$'. There is a slight danger of confusion—see what we say about functions of a complex variable—but if we take care it will be all right.*

7.2 Limits at infinity: functions of a complex variable

Definition 7.2. *Now let $E \subseteq \mathbb{C}$ and $f : E \to \mathbb{R}$ (or $\mathbb{C}$) and let $l \in \mathbb{R}$ (or $\mathbb{C}$). Suppose that for every $b \in \mathbb{R}$ there are points $z \in E$ such that $|z| > b$. We say that $f(z) \to l$ **as** $z \to \infty$ if, $\forall \varepsilon > 0$, $\exists B > 0$ such that $\forall z \in E$ such that $|z| > B$*

$$|f(z) - l| < \varepsilon.$$

We write this as $\lim_{z\to\infty} f(z) = l$.

Note that there may be a mild inconsistency with the previous definition if $E \subseteq \mathbb{R}$. If we are thinking 'complex' we'll need both the real limits at $\pm\infty$ to be equal.

Example 7.1. *Consider $\dfrac{\sin z}{z}$ as $z \to \infty$. For real values $z = x$ we get that $\left|\dfrac{\sin x}{x}\right| \leqslant \dfrac{1}{|x|} \to 0$ as $x \to \infty$. But for pure imaginary values like $z_k = 2\pi \mathrm{i} k$, with $k \in \mathbb{Z}$ we'll get that $\left|\dfrac{\sin z_k}{z_k}\right| = \dfrac{e^{2\pi k} - e^{-2\pi k}}{4\pi k} \to \infty$ as $k \to \infty$.*

Exercise 7.2. *Write down the contrapositive of 'f tends to a limit as $z \to \infty$.*

7.3 Tending to infinity...

Very briefly we discuss 'infinite limits'. We must take great care not to deceive ourselves: in neither $\mathbb{R}$ nor $\mathbb{C}$ is there a number ∞.

Definition 7.3. *Let $E \subseteq \mathbb{R}$ (or $\mathbb{C}$) and $f : E \to \mathbb{R}$ and let p be a limit point of E. We say that $f(z)$ **tends to** $+\infty$ **as** $z \to p$ if $\forall B > 0$, $\exists \delta > 0$ such that $\forall z \in E$ such that $0 < |z - p| < \delta$*

$$f(z) > B.$$

We may write this as $f(z) \to +\infty$ as $z \to p$.

Exercise 7.3. *Make a similar definition for $f(z) \to -\infty$ as $z \to p$.*

For complex valued functions things are easier:

Definition 7.4. *Let $E \subseteq \mathbb{R}$ (or $\mathbb{C}$) and $f : E \to \mathbb{C}$ and let p be a limit point of E. We say that $f(z)$* ***tends to*** ∞ ***as*** *$z \to p$ if $\forall B > 0\ \exists \delta > 0$ such that $\forall z \in E$ such that $0 < |z - p| < \delta$*

$$|f(z)| > B.$$

We may write this as $f(z) \to \infty$ as $z \to p$.

7.4 Euler's Limit

We prove the following result.

Proposition 7.2. *The limits $\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x$ and $\lim_{x\to-\infty} \left(1 + \frac{1}{x}\right)^x$ exist and are both equal to* e.

There are a number of ways of proving these. One method uses integration and looks at the area under the curve $1/x$. Another uses L'Hôpital's rule and will be given in Section 14. The following is a direct proof.

Proof. First limit: Recall that $\left(1 + \frac{1}{x}\right)^x := \exp\left(x \log\left(1 + \frac{1}{x}\right)\right)$. By the continuity of exp, from Problem Sheet 4,Q4b it is enough to prove that $\lim_{x\to\infty} x \log\left(1 + \frac{1}{x}\right) = 1$, or by AOL that $\lim_{x\to\infty} \frac{1}{x \log\left(1+\frac{1}{x}\right)} = 1$. Write $y = \log\left(1 + \frac{1}{x}\right)$; then

$$\frac{1}{x \log\left(1 + \frac{1}{x}\right)} - 1 = \frac{\exp(y) - 1 - y}{y}.$$

Note that as $1 + \frac{1}{x} > 1$ for $x > 0$, we have $y > 0$, and then

$$0 \leqslant \frac{\exp(y) - 1 - y}{y} = \frac{\sum_{n\geqslant 2} y^n/n!}{y} \leqslant \frac{\sum_{n\geqslant 2} y^n}{y} = \frac{y}{1 - y}.$$

So if we can show that $y \to 0$ as $x \to \infty$ we are done. But as log is continuous at 1 we can again use Problem Sheet 4,Q4b to see immediately that $y = \log(1 + \frac{1}{x}) \to 0$ as $x \to \infty$ as required.

A similar argument will deal with the other limit.

$\square$

8 Uniform Convergence

8.1 Motivation

Let $E \subseteq \mathbb{R}$ (or $\mathbb{C}$), and let $p \in E$ be a limit point, so that $p = \lim_{x\to p} x$. We have seen that 'continuity at p' is exactly the right condition to ensure that

$$\lim_{x\to p} f(x) = f(\lim_{x\to p} x),$$

that is to ensure that 'taking the limit $\lim_{x\to p}$' and 'finding the value under f' can be interchanged.

There are many other situations in which we would like to understand whether the *order* in which we perform two mathematical operations is significant or not:

(i) Suppose we have not just a single function f on E but a whole sequence (f_n). When is $\lim_{n\to\infty} \lim_{x\to p} f_n(x) = \lim_{x\to p} \lim_{n\to\infty} f_n(x)$?
In particular, if $f_n(x)$ is continuous at p, when is $\lim_{n\to\infty} f_n(x)$ continuous at p?

(ii) Similarly, when is $\lim_{x\to p} \sum_0^\infty f_n(x) = \sum_0^\infty \lim_{x\to p} f_n(x)$ and in particular if $f_n(x)$ is continuous at p when is $\sum_0^\infty f_n(x)$ continuous at p?

(iii) Once we have defined derivatives and integrals—as limits—we will want to know when $\lim_{n\to\infty} f_n'(x) = (\lim_{n\to\infty} f_n(x))'$? and when $\lim_{n\to\infty} \int_a^b f_n(t)\,\mathrm{d}t = \int_a^b \lim_{n\to\infty} f_n(t)\,\mathrm{d}t$? So when can we differentiate a series term by term and when can we integrate it term by term?

The answers to some of these questions are given in this lecture and the next.

To see that there are non-trivial problems we look at one typical example.

Example 8.1. *Consider the sequence of functions (f_n), where $f_n : [0,1] \to \mathbb{R}$ given by*

$$f_n(x) = \begin{cases} -nx+1 & \text{if } 0 \leqslant x < \frac{1}{n}, \\ 0 & \text{if } x \geqslant \frac{1}{n}. \end{cases}$$

Consider also the function $f : [0,1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{if } x > 0. \end{cases}$$

Sketch their graphs, and note that for all $x \in [0,1]$ we have that $f(x) = \lim_{n\to\infty} f_n(x)$.

Note that although all the f_n are continuous the limit function f is not continuous at 0.

[*Once we have Theorem 8.2 we will be able to prove that f_n is not uniformly convergent to f on $[0,1]$. Hint: Consider $x_n = 1/(2n)$. Then $f_n(x_n) = 1/2$ so that $\sup_{x\in[0,1]} |f_n(x) - f(x)| \geq f_n(x_n) = 1/2 \not\to 0$. Alternatively once we have Theorem 9.1 we can say immediately that the convergence is not uniform as we have a sequence of continuous functions whose limit is not continuous.*]

8.2 Definition

As the sum of an infinite series is defined as the limit of the sequence of partial sums we will start by looking at sequences of functions.

Let $E \subseteq \mathbb{R}$ (or $\mathbb{C}$) and let $f_n : E \to \mathbb{R}$ (or $\mathbb{C}$) be a sequence of functions. Then for each (fixed) $x \in E$, $(f_n(x))$ is a sequence of real (or complex) numbers. If this sequence converges for every $x \in E$, then the limit which will depend on x so we will call it $f(x)$. Thus $f : E \to \mathbb{R}$ (or $\mathbb{C}$) is a function. Hence we have the definition (using Analysis I):

Definition 8.1. *By f_n* ***converges to** f* ***on** E we mean that $\forall x \in E$, and $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n > N$*

$$|f_n(x) - f(x)| < \varepsilon.$$

So, of course, in general N depends on x.

Just as when we defined 'uniform continuity' as a stronger version of 'continuous at all points' by insisting on being able to choose one 'δ' to deal with all points, so we now strengthen our definition of 'convergence of a sequence of functions'. For 'uniform convergence' we insist that one N works for all x.

Definition 8.2. *By f_n* ***converges uniformly to** f* ***on** E we mean that $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n > N$ and $\forall x \in E$*

$$|f_n(x) - f(x)| < \varepsilon.$$

We write this as '$f_n \to f$ uniformly on E' or '$f_n \xrightarrow{u} f$'.

It is trivial to see that:

Proposition 8.1. *If the sequence (f_n) converges uniformly to f on E then at every point $x \in E$ we have that the sequence $(f_n(x))$ converges to $f(x)$.*

There is one special case which we should single out. Suppose that for each $n \in \mathbb{N}$ we have that $s_n(x) = \sum_0^n f_k(x)$ and that $s : E \to \mathbb{R}$ (or $\mathbb{C}$). If we apply the definition to the sequence (s_n) and the function s we will get

Definition 8.3. *We say that* ***the series** $\sum f_n$* ***converges uniformly to** s* ***on** E if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n > N$ and $\forall x \in E$*

$$\left| \sum_{k=0}^{n} f_k(x) - s(x) \right| < \varepsilon.$$

We may write this as '$\sum_0^\infty f_n(x) = s(x)$ (uniformly on E)'.

8.3 Test for Uniform Convergence

We can re-express the definition in a more practical way:

Theorem 8.2. *Let E be a non-empty subset of $\mathbb{R}$ or $\mathbb{C}$. Let $f_n, f : E \to \mathbb{R}$ (or $\mathbb{C}$). Then the following are equivalent:*

(i) $f_n \to f$ *uniformly on E;*

(ii) $\exists N$ *s.t.* $\forall n > N$, $m_n := \sup_{x \in E} |f_n(x) - f(x)|$ *exists and* $m_n \to 0$ *as* $n \to \infty$.

Proof. ($\Longrightarrow$)

Suppose $f_n \to f$ uniformly on E. That is $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall x \in E$ and $\forall n > N$

$$|f_n(x) - f(x)| < \tfrac{1}{2}\varepsilon.$$

Hence, for each $n > N$, $\frac{1}{2}\varepsilon$ is an upper bound of the set $\{|f_n(x) - f(x)| : x \in E\}$. So m_n exists and

$$m_n = \sup_{x \in E} |f_n(x) - f(x)| \leqslant \tfrac{1}{2}\varepsilon < \varepsilon \quad \forall n > N.$$

By the definition of sequence limits, $\lim_{n\to\infty} m_n = 0$.

($\Longleftarrow$)

Suppose the m_n exist for all $n > N_1$, and that $\lim_{n\to\infty} \sup_{x \in E} |f_n(x) - f(x)| = 0$, Then $\forall \varepsilon > 0$ $\exists N > N_1$ such that $\forall n > N$

$$\sup_{x \in E} |f_n(x) - f(x)| < \varepsilon.$$

Therefore

$$|f_n(x) - f(x)| \leqslant \sup_{x \in E} |f_n(x) - f(x)| < \varepsilon \quad \forall x \in E \text{ and } \forall n > N.$$

That is $f_n \to f$ *uniformly on* E. □

Example 8.2. *Let* $E = [0, 1)$ *and let* $f_n(x) = x^n$. *Clearly* $\lim_{n\to\infty} f_n(x) = 0$, *so* $f(x) = 0$. *Then* $m_n = \sup_{x\in E} |x^n - 0| = \sup_{x \in E} x^n$. *But* $x_n = (1/2)^{\frac{1}{n}} \in E$ *and* $f_n(x_n) = 1/2$ *so that*

$$m_n \geqslant f_n(x_n) = 1/2 \nrightarrow 0, \qquad \textit{as } n \to \infty$$

so f_n *is not uniformly convergent on* $[0, 1)$.

However, if instead we consider $E = [0, r]$, *where* $0 < r < 1$ *is a* fixed *constant. Then* $x^n \to 0$ *uniformly on* E, *because now*

$$m_n = \sup_{[0,r]} x^n \leqslant r^n \to 0, \qquad \textit{as } n \to \infty.$$

Remark 8.3. *The test is particularly useful if* $E = [a, b]$ *and the functions* f_n *and* f *are differentiable. In such cases the supremum will be achieved either at* a *or at* b *or at some interior point where* $\dfrac{\mathrm{d}(f_n(x) - f(x))}{\mathrm{d}x} = 0$. *We will prove this later in the course; for the moment you can use it in exercises.*[7]

[7] Of course we will not use it in building up the theory.

8.4 Cauchy's Criterion

Just as we found for sequences of numbers there is a characterisation of uniform convergence which does not depend on knowing the limit function.

Theorem 8.4 (Cauchy's Criterion for Uniform Convergence). *Let $E \subseteq \mathbb{R}$ (or $\mathbb{C}$) and let $f_n : E \to \mathbb{R}$ (or $\mathbb{C}$). Then f_n converges uniformly on E, if and only if, $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n, m > N$ and $\forall x \in E$*

$$|f_n(x) - f_m(x)| < \varepsilon. \qquad (*)$$

Proof. ($\Longrightarrow$) Suppose f_n converges uniformly on E with limit function f, then $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n > N$ and $\forall x \in E$

$$|f_n(x) - f(x)| < \tfrac{1}{2}\varepsilon.$$

So, $\forall x \in E$ and $\forall n, m > N$

$$\begin{aligned} |f_n(x) - f_m(x)| &\leqslant |f_n(x) - f(x)| + |f_m(x) - f(x)| \\ &< \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon \\ &= \varepsilon. \end{aligned}$$

($\Longleftarrow$) Conversely, suppose (*) holds. Then for any $x \in E$, $(f_n(x))$ is a Cauchy sequence, so that it is convergent. Let us denote its limit by $f(x)$. For every $\varepsilon > 0$, choose $N \in \mathbb{N}$ such that $\forall n, m > N$ and $\forall x \in E$

$$|f_n(x) - f_m(x)| < \tfrac{1}{2}\varepsilon.$$

Now fix $n > N$ and $x \in E$, and let $m \to \infty$ in the above inequality. So, by the preservation of weak inequalities

$$|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leqslant \tfrac{1}{2}\varepsilon < \varepsilon.$$

Hence $f_n \to f$ uniformly on E. □

Corollary 8.5 (Cauchy's criterion for uniform convergence of series)**.** *The series $\sum_{n=0}^{\infty} f_n$ is uniformly convergent on E if and only if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n > m > N$ and $\forall x \in E$*

$$\left| \sum_{k=m+1}^{n} f_k(x) \right| < \varepsilon.$$

8.5 The M-test

As a consequence, we prove the following simple but very important test for uniform convergence of series.

Theorem 8.6 (The Weierstrass M-Test). *Let $E \subseteq \mathbb{R}$ (or $\mathbb{C}$) and $f_n : E \to \mathbb{R}$ (or $\mathbb{C}$). Suppose that there is a sequence (M_n) of real numbers such that*

$$|f_n(x)| \leqslant M_n \qquad \forall x \in E.$$

If $\sum_{n=0}^{\infty} M_n$ converges then $\sum_{n=0}^{\infty} f_n$ converges uniformly on E.

Note that the M_n must be independent of x.

Proof. By Cauchy's Criterion for the convergence of $\sum M_n$ we have that $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n > m > N$

$$\sum_{k=m+1}^{n} M_k < \varepsilon.$$

Now by the Triangle Law

$$\left| \sum_{k=m+1}^{n} f_k(x) \right| \leqslant \sum_{k=m+1}^{n} |f_k(x)| \leqslant \sum_{k=m+1}^{n} M_k < \epsilon, \forall n > m > N \text{ and } \forall x \in E,$$

which is Cauchy's criterion for the uniform convergence of the series. □

Corollary 8.7. *Suppose the conditions for the M-test hold, and $\sum M_n$ is convergent. Then*

$$\left| \sum_{n=0}^{\infty} f_n(x) \right| \leqslant \sum_{n=0}^{\infty} |f_n(x)| \leqslant \sum_{n=0}^{\infty} M_n \quad \forall x \in E.$$

Proof. Apply the preservation of weak inequalities as $N \to \infty$ to the obvious inequalities

$$\left| \sum_{n=0}^{N} f_n(x) \right| \leqslant \sum_{n=0}^{N} |f_n(x)| \leqslant \sum_{n=0}^{N} M_n \quad \forall x \in E.$$

□

9 Uniform Convergence: Examples and Applications

9.1 Examples

Example 9.1. *Let $E = [0, 1]$ and let*

$$f_n(x) = \frac{nx}{1 + n^2x^2}.$$

Then clearly $\lim_{n\to\infty} f_n(x) = 0$ for every $x \in [0, 1]$.

But $f_n(1/n) = 1/2$*, so that*

$$\sup_{x\in[0,1]} |f_n(x) - f(x)| \geqslant \frac{1}{2} \nrightarrow 0 \text{ as } n \to \infty$$

and so f_n *converges to* 0 *but* ***not uniformly*** *in* $[0, 1]$.

Example 9.2. $\sum_{n=0}^{\infty} x^n$ *converges to* $\frac{1}{1-x}$ *in* $(-1, 1)$*, but not uniformly.*

From Analysis I, $s_n(x) = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$ *tends to* $\frac{1}{1-x}$ *for any* $|x| < 1$*. On the other hand*

$$\left| s_n(x) - \frac{1}{1-x} \right| = \frac{|x|^{n+1}}{|1-x|}$$

so that (look at $x = \frac{n+1}{n+2}$*)*

$$\sup_{x\in(-1,1)} \left| s_n(x) - \frac{1}{1-x} \right| \geqslant \frac{\left(\frac{n+1}{n+2}\right)^{n+1}}{|1 - \frac{n+1}{n+2}|} = \frac{n+2}{\left(1 + \frac{1}{n+1}\right)^{n+1}} \to \infty.$$

Hence $\sum_{n=0}^{\infty} x^n$ *doesn't converge uniformly.*

Example 9.3. $\sum_{n=0}^{\infty} x^n$ *converges uniformly on* $[-r, r]$ *for any* $0 < r < 1$.

This follows from the M*-test with* $M_n := r^n$.

9.2 Uniform Convergence preserves continuity

We have already seen that the limit of a sequence of continuous functions may not be continuous. This very important theorem tells us that 'uniformity' gives us the extra condition we need.

Theorem 9.1. *Let* f_n,$f : E \to \mathbb{R}$ *(or* $\mathbb{C}$*), and* $f_n \to f$ *uniformly in* E*. Suppose all* f_n *are continuous at* $x_0 \in E$*. Then the limit function* f *is also continuous at* x_0*, so that*

$$\lim_{x\to x_0} \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} f_n(x_0) = \lim_{n\to\infty} \lim_{x\to x_0} f_n(x).$$

Proof. $\forall\varepsilon > 0\ \exists N \in \mathbb{N}$ s.t. $\forall n > N$ and $\forall x \in E$

$$|f_n(x) - f(x)| < \tfrac{1}{3}\varepsilon.$$

Since f_{N+1} is continuous at x_0, $\exists\delta > 0$ (depending on x_0 and ε) such that $\forall x \in E$ such that $|x - x_0| < \delta$

$$|f_{N+1}(x) - f_{N+1}(x_0)| < \tfrac{1}{3}\varepsilon.$$

Hence, if $x \in E$ and $|x - x_0| < \delta$, then by the Triangle Law

$$\begin{aligned}
& |f(x) - f(x_0)| \\
\leqslant\ & |f(x) - f_{N+1}(x)| + |f_{N+1}(x) - f_{N+1}(x_0)| + |f_{N+1}(x_0) - f(x_0)| \\
<\ & \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon \\
=\ & \varepsilon.
\end{aligned}$$

By definition, f is continuous at x_0. □

Note it is very important that $N+1$ is fixed, so that δ does not depend on n.

Remark 9.2 (Version for series). *If $\sum_{n=0}^{\infty} f_n$ converges uniformly on E and every f_n is continuous at $x_0 \in E$, then the function $\sum_{n=0}^{\infty} f_n(x)$ is continuous at x_0, that is*

$$\lim_{x \to x_0} \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} f_n(x_0).$$

In particular, if f_n is continuous on E for all n and $\sum_{n=0}^{\infty} f_n$ converges uniformly on E, then $\sum_{n=0}^{\infty} f_n$ is continuous on E.

9.3 Power Series

We can apply the the results of the previous subsection to the important case of power series.

Theorem 9.3 (Continuity of Power Series). *Suppose the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n x^n$ is R, where $0 \leqslant R \leqslant \infty$. Then for every $0 \leqslant r < R$, $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on the closed disk $\{x : |x| \leqslant r\}$. Therefore, $\sum_{n=0}^{\infty} a_n x^n$ is continuous on the open disk $\{x : |x| < R\}$.*

Proof. By the definition of 'radius of convergence', $\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent for $|x| < R$. In particular, $\sum_{n=0}^{\infty} |a_n| r^n$ is convergent. Since

$$|a_n x^n| \leqslant |a_n| r^n \quad \text{for all } x \quad \text{such that } |x| \leqslant r$$

we have, by Weierstrass M-test with $M_n = |a_n| r^n$, that $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $\{x : |x| \leqslant r\}$.

But $a_n x^n$ is continuous for any $n \in \mathbb{N}$. So, for any $r < R$, $\sum_{n=0}^{\infty} a_n x^n$ is continuous for $|x| \leqslant r$, and hence on the open disk $\{x : |x| < R\}$. □

Example 9.4. *Note that in general it is not true that the power series is uniformly convergent on $|x| < R$. For example if a power series $\sum_{n=0}^{\infty} a_n x^n$ is uniformly convergent on $\mathbb{R}$, then there exists $N \in \mathbb{N}$ such that $a_n = 0$ for all $n > N$.*

Proof: By CC for uniform convergence $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that for all $n > m > N$ and $\forall x \in \mathbb{R}$, $|\sum_{r=m+1}^{n} a_r x^r| < \epsilon$. So in particular for all $r > N+1$ and all $x \in \mathbb{R}$

$$|a_r x^r| < \epsilon.$$

Hence $a_r = 0$ for all $r > N+1$.

Note 9.4. *Note that Theorem 9.3 says nothing about convergence or continuity at the end-points. If you are interested, subsection 9.5 deals with this in the real case.*

Corollary 9.5. *The functions $\exp x$, $\sin x$, $\cos x$, $\cosh x$ and $\sinh x$ can all be defined by power series with infinite radius of convergence so are all continuous on $\mathbb{C}$.*

9.4 Integrals and derivatives of sequences

Next term, in the course Analysis III, you will learn how to define integrals, and the proofs of the following theorems will be given.

Theorem 9.6. *If $f_n \to f$ uniformly on $[a, b]$ and if every f_n is continuous, then*

$$\int_a^b f = \int_a^b \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_a^b f_n.$$

Similarly, if the series $\sum_{n=1}^{\infty} f_n$ converges uniformly on $[a, b]$ and if all f_n are continuous, then we may integrate the series term by term

$$\int_a^b \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int_a^b f_n.$$

Note 9.7. *However, uniform convergence is not the 'right' condition for integrating a series term by term: we can exchange the order of integration $\int_a^b$ (which involves a limiting procedure) and $\lim_{n\to\infty}$ under much weaker conditions. The search for correct conditions for term-by-term integration led to the discovery of Lebesgue integration [Part A option: Integration].*

Theorem 9.8. *Let $f_n(x) \to f(x)$ for each $x \in [a, b]$. Suppose f_n' exists and is continuous on $[a, b]$ for every n, and that $f_n' \to g$ uniformly on $[a, b]$. Then f' exists and is continuous on $[a, b]$, and $\forall x \in [a, b]$*

$$\frac{\mathrm{d}}{\mathrm{d}x} \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_n(x).$$

Similarly, if $\sum_{n=0}^{\infty} f_n$ converges on $[a, b]$, and if every f_n' exists and is continuous on $[a, b]$ and if $\sum_{n=0}^{\infty} f_n'$ converges uniformly on $[a, b]$, then $\sum_{n=0}^{\infty} f_n$ is differentiable on $[a, b]$, its derivative is continuous, and $\forall x \in [a, b]$

$$\frac{\mathrm{d}}{\mathrm{d}x} \sum_{n=0}^{\infty} f_n(x) = \sum_{n=01}^{\infty} f_n'(x).$$

9.5 The end points

This section will be omitted from lectures and is included for interest.

When $0 < R < \infty$ the points where $|z| = R$ need to be handled differently. We only deal with the real case, so there are two such points R and $-R$. Scaling (replacing x by x/R or $-x/R$) lets us deal only with power series where the radius is 1 and describe what happens at $x = 1$.

Theorem 9.9 (Abel's Continuity Theorem). *Suppose that the series* $\sum_{n=0}^{\infty} a_n x^n$ *has radius of convergence* $R = 1$. *Suppose further that* $\sum_{n=0}^{\infty} a_n$ *converges.*

Then $\sum_{n=0}^{\infty} a_n x^n$ *converges uniformly on* $[0, 1]$.

Consequently, $\sum_{n=0}^{\infty} a_n x^n$ *is continuous on* $(-1, 1]$, *and in particular*

$$\lim_{x \to 1-} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n.$$

Proof. First note that our general result gives continuity on $(-1, 1)$; it is only the point $x = 1$ we have to deal with. We will get continuity provided we get uniform convergence on $[0, 1]$.

By Cauchy's Criterion for the convergent $\sum_{n=0}^{\infty} a_n$ we have that, for every $\varepsilon > 0$, there is N such that, for every $n > m > N$ we have

$$\left| \sum_{k=m}^{n} a_k \right| < \varepsilon.$$

Now fix $m > N$, and for the partial sums from m use the notation

$$A_k = \sum_{j=m}^{k} a_j \quad \text{for } k \geqslant m; \quad \text{and } A_{m-1} = 0$$

noting that subtracting consecutive sums gives us back the original sequence[8]

$$a_k = A_k - A_{k-1}.$$

By what we have from the Cauchy Criterion above, $|A_k| < \varepsilon$ whenever $k \geqslant m - 1$. We have by elementary algebra the following formula[9]

$$\begin{aligned} \sum_{k=m}^{n} a_k x^k &= \sum_{k=m}^{n} (A_k - A_{k-1}) x^k \\ &= \sum_{k=m}^{n} A_k x^k - \sum_{k=m}^{n} A_{k-1} x^k \\ &= \sum_{k=m}^{n-1} A_k \left(x^k - x^{k+1} \right) + A_n x^n. \end{aligned}$$

Hence, by the Triangle Law we have that

$$\begin{aligned} \left| \sum_{k=m}^{n} a_k x^k \right| &\leqslant \sum_{k=m}^{n-1} |A_k| \left(x^k - x^{k+1} \right) + |A_n| x^n \\ &< \varepsilon \sum_{k=m}^{n-1} \left(x^k - x^{k+1} \right) + \varepsilon x^n \\ &= \varepsilon x^m \\ &\leqslant \varepsilon \end{aligned}$$

[8] Think 'Differentiation undoes Integration'.

[9] This is called Abel's summation formula—think 'integration by parts'.

for any $x \in [0,1]$.

The Cauchy Criterion yields that $\sum_{n=0}^{\infty} a_n x^n$ is uniformly convergent on $[0,1]$. $\square$

9.6 Monotone Sequences of Continuous Functions

This section will be omitted from lectures and is included for interest.

The theorem of this subsection is a partial converse of our theorem that 'uniform convergence preserves continuity'; if the sequence is monotone then the continuity of the limit will give uniformity of convergence.

Theorem 9.10 (**The Dini Theorem**)**.** . *Let f_n be a sequence of real continuous functions on $[a,b]$; and let f be a real continuous function on $[a,b]$.*

Suppose that

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad \text{for every } x \in [a,b]$$

and that

$$f_n(x) \geqslant f_{n+1}(x) \qquad \text{for all } \ n \quad \text{and for all } x \in [a,b].$$

Then $f_n \to f$ uniformly on $[a,b]$.

Proof. Let $g_n(x) = f_n(x) - f(x)$. Then g_n is continuous for every n, $g_n \geqslant 0$ and $\lim_{n\to\infty} g_n(x) = 0$ for any $x \in [a,b]$. Suppose (g_n) were not uniformly convergent on $[a,b]$. Write down the contrapositive to see that for some $\varepsilon > 0$, and every natural number k there exists a natural number $n_k > k$ and a point $x_k \in [a,b]$ such that

$$|g_{n_k}(x_k)| = g_{n_k}(x_k) \geqslant \varepsilon.$$

We may choose n_k so that $k \to n_k$ is increasing. We may assume that $x_k \to p$—otherwise use the Bolzano–Weierstrass theorem to extract a convergent subsequence of (x_k) and use it instead. Then $p \in [a,b]$. For any (fixed) k, since (g_n) is decreasing,

$$\varepsilon \leqslant g_{n_l}(x_l) \leqslant g_{n_k}(x_l)$$

for all $l > k$. Letting $l \to \infty$ in the above inequality, we obtain

$$\varepsilon \leqslant \lim_{l\to\infty} g_{n_k}(x_l) = g_{n_k}(p)$$

as g_{n_k} is continuous at p. This contradicts to the assumption that $\lim_{k\to\infty} g_{n_k}(p) = 0$. $\square$

Example 9.5. *Let $f_n(x) = \frac{1}{1+nx}$ for $x \in (0,1)$. Then $\lim_{n\to\infty} f_n(x) = 0$ for every $x \in (0,1)$, f_n is decreasing in n, but f_n does not converge uniformly. Dini's theorem doesn't apply, as $(0,1)$ is not compact.*

10 Differentiation: definitions and elementary results

10.1 Definitions

In this course we only study differentiability for real (or complex)-valued functions on E, where E is a subset of the real line $\mathbb{R}$. The theory of the differentiability of complex valued functions on the complex plane $\mathbb{C}$ is very different from the real case and requires another theory—See Complex Analysis [Part A: Analysis]. Generally E will be an interval.

Definition 10.1. *Let $f : (a,b) \to \mathbb{R}$ (or $\mathbb{C}$), and let $x_0 \in (a,b)$. By f **is differentiable at** x_0 we mean that the following limit exists:*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

*When it exists we denote the limit by $f'(x_0)$ which we call the **derivative of** f **at** x_0.*

[That is $\forall \varepsilon > 0\ \exists \delta > 0$ such that $\forall x \in (a,b)$ such that $0 < |x - x_0| < \delta$

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < \varepsilon.]$$

For example, it is easy to see that the function $f(x) = x$ is differentiable at every point of $\mathbb{R}$ and has derivative $f'(x_0) = 1$ at every point; and the function $g(t) = \mathrm{e}^{2\pi \mathrm{i} t}$ is differentiable at every point, although we can't yet prove that.

Sometimes it is helpful to also define 'left-hand' and 'right-hand' versions of these.

Definition 10.2. *(i) Let $f : [a,b) \to \mathbb{R}$ (or $\mathbb{C}$), and let $x_0 \in [a,b)$. We say that f **has a right-derivative** at x_0 if the following limit exists*

$$\lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0}.$$

If the limit exists we denote it by $f'_+(x_0)$.

*(ii) Let $f : (a,b] \to \mathbb{R}$ (or $\mathbb{C}$), and let $x_0 \in (a,b]$. We say that f **has a left-derivative** at x_0 if the following limit exists*

$$\lim_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}.$$

If the limit exists we denote it by $f'_-(x_0)$.

The following result is easily proved (compare what we did for left- and right-continuity).

Proposition 10.1. *Let $f : (a,b) \to \mathbb{R}$ (or $\mathbb{C}$). Then the following are equivalent:*

(a) f is differentiable at x_0 and $f'(x_0) = l$;

(b) f has both left- and right-derivatives at x_0, and $f'_-(x_0) = l = f'_+(x_0)$.

Definition 10.3. *(i) Suppose that $f : (a,b) \to \mathbb{R}$ (or $\mathbb{C}$). Then we say that f **is differentiable on** (a,b) if f is differentiable at every point of (a,b).*

*(ii) Suppose that $f : [a,b] \to \mathbb{R}$ (or $\mathbb{C}$). Then we say that f **is differentiable on** $[a,b]$ if f is differentiable at every point of (a,b), and if $f'_+(a)$ and $f'_-(b)$ exist.*

If you wish you can define differentiable on $(a,b]$ and $[a,b)$ as well.

Remark 10.2. *Let $y = f(x)$. There are other notations for derivatives*

$\frac{\mathrm{d}y}{\mathrm{d}x}$ *or* $\frac{\mathrm{d}f(x_0)}{\mathrm{d}x}$ *[G. W. Leibnitz]*

y' *or* $f'(x_0)$ *[J. L. Lagrange]*

Dy *or* $Df(x_0)$ *[A. L. Cauchy, in particular for vector-valued functions of several variables].*

10.2 An Example

Define a function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{for } x > 0, \\ 0 & \text{for } x \leqslant 0. \end{cases}$$

Then we can show that

$$f'(x) = \begin{cases} 0 & \text{when } x < 0, \\ 0 & \text{when } x = 0, \\ 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x} & \text{when } x > 0. \end{cases}$$

The derivative for $x \leqslant 0$ can be found directly from the definition. Later we will see that we can use the chain rule to find the derivative for $x > 0$.

Note that the derivative is not continuous at the origin. (See problem sheet 5.)

We can get other interesting examples by replacing the 'x^2 by x^α and the '$\dfrac{1}{x}$' by $\dfrac{1}{x^\beta}$.

10.3 Derivatives and differentials

By looking at the definition of 'limit' in terms of ε and δ (see problem sheet) we can easily prove that:

Proposition 10.3. *Suppose that $f : (a, b) \to \mathbb{R}$ is differentiable at $x_0 \in (a, b)$ and that $f'(x_0) > 0$. Then there exists a $\delta > 0$ such that for all $x \in (x_0, x_0 + \delta)$ we have that $f(x) > f(x_0)$, and for all $x \in (x_0 - \delta, x_0)$ we have that $f(x) < f(x_0)$.*

We have corollaries like:

Corollary 10.4. *Suppose that $f : [a, b) \to \mathbb{R}$ is right-differentiable at $x_0 \in [a, b)$ and that $f'_+(x_0) > 0$. Then there exists a $\delta > 0$ such that for all $x \in (x_0, x_0 + \delta)$ we have that $f(x) > f(x_0)$.*

In fact, if f is differentiable at x_0, then the 'increment' of f near x_0 can be expressed

$$f(x) - f(x_0) = f'(x_0)\,(x - x_0) + o(x - x_0)$$

where o is a function of x and x_0 satisfying

$$\lim_{x \to x_0} \frac{o(x - x_0)}{x - x_0} = 0.$$

That is, the 'linear part' of the increment $f(x) - f(x_0)$ is $f'(x_0)\,(x - x_0)$; all the rest is small in comparison. This is sometimes called *the differential* of f at x_0. It is the first approximation to f near x_0.

10.4 Differentiability and Continuity

Theorem 10.5 (Differentiability $\Longrightarrow$ Continuity). *Let $f : (a, b) \to \mathbb{R}$ (or $\mathbb{C}$). If f is differentiable at $x_0 \in (a, b)$ then f is continuous at x_0.*

Proof. Since

$$\begin{aligned}
\lim_{x \to x_0} (f(x) - f(x_0)) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}\,(x - x_0) \\
&= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \lim_{x \to x_0} (x - x_0) \text{ by AOL} \\
&= f'(x_0) \times 0 \\
&= 0.
\end{aligned}$$

Therefore $\lim_{x \to x_0} f(x) = f(x_0)$, so that, by definition, f is continuous at x_0. ☐

Note: The converse is not true. For example $|x|$ is continuous but is not differentiable at 0. In fact there exist functions which are continuous everywhere, but not differentiable at any point! (See Bartle and Sherbert.)

10.5 Algebraic properties

The following results are straightforward consequences of the Algebra of Limits. They let us build up at once all the calculus we learned at school—once we can differentiate a few standard functions (constants, linear functions, exp, sin and cos).

Theorem 10.6. *Suppose $f, g : (a, b) \to \mathbb{R}$ (or $\mathbb{C}$) are both differentiable at $x_0 \in (a, b)$, and $\lambda, \mu \in \mathbb{R}$ (or $\mathbb{C}$).*
(i) [Linearity of differentiation] $\lambda \cdot f + \mu \cdot g$ is differentiable at x_0 and

$$(\lambda \cdot f + \mu \cdot g)'(x_0) = \lambda \cdot f'(x_0) + \mu \cdot g'(x_0).$$

(ii)[The Product Rule] $fg : x \mapsto f(x)g(x)$ is differentiable at x_0 and

$$(fg)'(x_0) = f(x_0)g'(x_0) + f'(x_0)g(x_0).$$

(iii) [The Quotient Rule] Suppose $g(x_0) \neq 0$. Then $x \mapsto \frac{f(x)}{g(x)}$ is differentiable at x_0 and

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

Proof. (ii) Apply AOL to

$$\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = f(x)\frac{g(x) - g(x_0)}{x - x_0} + g(x_0)\frac{f(x) - f(x_0)}{x - x_0}.$$

Let $x \to x_0$ and use the definitions of $f'(x_0)$, $g'(x_0)$, and the continuity of $f(x)$ so $f(x) \to f(x_0)$.

(iii)See problem sheet 5.

□

So now we know that x^n is differentiable at all points, so are polynomials, and also rational functions at points where the denominator is non zero.

10.6 The Chain Rule

Theorem 10.7 (The Chain Rule)**.** *Suppose $f : (a, b) \to \mathbb{R}$, and that $g : (c, d) \to \mathbb{R}$. Suppose that $f((a, b)) \subseteq (c, d)$, so that $g \circ f : (a, b) \to \mathbb{R}$ is defined.*

Suppose further that f is differentiable at $x_0 \in (a, b)$, and that g is differentiable at $f(x_0)$.

Then $g \circ f$ is differentiable at x_0 and

$$(g \circ f)'(x_0) = g'(f(x_0))\, f'(x_0).$$

Proof. Write $y_0 = f(x_0)$, and define a function v on (c, d) by

$$v(y) = \begin{cases} \dfrac{g(y) - g(y_0)}{y - y_0} - g'(y_0) & \text{for all } \ y \neq y_0, \\ 0 & \text{for } \ y = y_0. \end{cases}$$

Note that $v(y) \to 0$ as $y \to y_0$, so that v is continuous at y_0.

Rearranging the definition of v we see that for all $y \in (c, d)$

$$g(y) - g(y_0) = (y - y_0)\,(g'(y_0) + v(y))$$

In particular

$$g(f(x)) - g(f(x_0)) = (f(x) - f(x_0))\,(g'(y_0) + v(f(x)))$$

so that

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = g'(y_0)\frac{f(x) - f(x_0)}{x - x_0} + v(f(x))\frac{f(x) - f(x_0)}{x - x_0}.$$

Since f is differentiable at x_0, f continuous at x_0. But v is continuous at $y_0 = f(x_0)$ and hence $v(f(x))$ is continuous at x_0. Thus $v(f(x)) \to 0$ as $x \to x_0$. Letting $x \to x_0$ we obtain, using AOL

$$\begin{aligned} \lim_{x \to x_0} \frac{g(f(x)) - g(f(x_0))}{x - x_0} &= g'(y_0) \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ &\quad + \lim_{x \to x_0} v(f(x)) \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ &= g'(y_0) f'(x_0) + 0 \times f'(x_0) \\ &= f'(x_0) g'(y_0)\ . \end{aligned}$$

□

10.7 Higher Derivatives

Suppose that $f : (a, b) \to \mathbb{R}$ (or $\mathbb{C}$) is differentiable at every point of some $(x_0 - \delta, x_0 + \delta)$. Then it makes sense to ask if f' is differentiable at x_0. If it is differentiable then we denote its derivative by $f''(x_0)$.

More generally we can define the $(n + 1)$-th derivative $f^{(n+1)}$ recursively.

Definition 10.4. *Suppose that $f : (a, b) \to \mathbb{R}$ (or $\mathbb{C}$) is such that f, f',...,$f^{(n)}$ exist at every point of (a, b). Suppose that $x_0 \in (a, b)$. By f* ***is*** $(n+1)$***-times differentiable at*** *x_0 we mean that $f^{(n)}$ is differentiable at x_0. We write $f^{(n+1)}(x_0) := f^{(n)\prime}(x_0)$.*

If f has derivatives of all orders on (a, b) we sometimes say it is **infinitely differentiable.**

The following is proved by an easy induction using Linearity and the Product Rule.

Theorem 10.8 (The Leibnitz Formula)**.** *Let $f, g : (a, b) \to \mathbb{R}$ (or $\mathbb{C}$) be n-times differentiable on (a, b). Then $x \mapsto f(x)g(x)$ is n-times differentiable and*

$$(fg)^{(n)}(x) = \sum_{j=0}^{n} \binom{n}{j} f^{(j)}(x) g^{(n-j)}(x).$$

11 The elementary functions

11.1 Differentiating power series

The elementary functions—$\exp x$, $\cos x$, $\sin x$, $\log x$, $\arctan x$—are defined as power series, or are got as inverse functions of real functions defined by power series.

We start with a lemma:

Lemma 11.1. *The power series $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=1}^{\infty} n a_n x^{n-1}$ have the same radius of convergence.*

Proof. Let the radii be R and R'; we will show $R \geqslant R'$ and $R' \geqslant R$.

$R \geqslant R'$: First suppose that $|x_1| < R'$; then $\sum_{n=1}^{\infty} n a_n x^{n-1}$ is absolutely convergent at $x = x_1$. That is, $\sum_{n=1}^{\infty} n|a_n||x_1|^{n-1}$ converges. Now note that $|a_n x_1^n| \leqslant n|a_n||x_1^n|$. Hence by the comparison test $\sum_{n=0}^{\infty} |a_n||x_1|^n$ converges.Therefore, by definition of 'radius of convergence' we have that $R \geqslant R'$.

$R' \geqslant R$: Now suppose that $|x_1| < R$; and choose x_2 so that $|x_1| < |x_2| < R$. Then $\sum_{n=0}^{\infty} |a_n||x_2|^n$ converges, and so (Analysis I) $|a_n||x_2|^n \to 0$ as $n \to \infty$. But a convergent sequence is bounded (Analysis I) so there exists M such that $|a_n||x_2|^{n-1} < M$ for all n. Now

$$n|a_n||x_1|^{n-1} \leqslant Mn \left| \frac{x_1}{x_2} \right|^{n-1}$$

and as, by the Ratio Test $\sum n \left| \frac{x_1}{x_2} \right|^{n-1}$ is convergent, we have by the Comparison Test that $\sum_{n=1}^{\infty} n|a_n||x_1|^{n-1}$ is convergent. By the definition of 'radius of convergence' we have that $R' \geqslant R$. □

Theorem 11.2 (Term-by-term differentiation)**.** *The power series $f(x) := \sum_{n=0}^{\infty} a_n x^n$ and $g(x) := \sum_{n=1}^{\infty} n a_n x^{n-1}$ have the same radius of convergence R, and for any x such that $|x| < R$ we have that f is differentiable at x and moreover that $f'(x) = g(x)$.*

Proof. (not examinable and will be omitted from the lectures)

The first part is done by the lemma.

Suppose $|x| < R$; choose some r such that $|x| < r < R$. (For example, $r := (|x| + R)/2$ if $R < \infty$, or $r = |x| + 1$ if $R = \infty$.)

For any point w such that $|w| < r$, consider

$$\begin{aligned} \frac{f(w) - f(x)}{w - x} - g(x) &= \sum_{n=1}^{\infty} a_n \left(\frac{w^n - x^n}{w - x} - n x^{n-1} \right) \\ &= \sum_{n=2}^{\infty} a_n \left(\frac{w^n - x^n}{w - x} - n x^{n-1} \right); \end{aligned}$$

where we have added the series $f(w)$, $f(x)$ and $g(x)$ term by term, which is justified by AOL. Our aim is to show that

$$\frac{f(w) - f(x)}{w - x} - g(x) \to 0 \qquad \text{as } w \to x.$$

The binomial identity

$$\frac{w^n - x^n}{w - x} = x^{n-1} + x^{n-2} w + \cdots + x w^{n-2} + w^{n-1}$$

is easily proved by induction; then we have that for any $w \neq x$ and $n \geqslant 2$

$$\begin{aligned} \frac{w^n - x^n}{w - x} - n x^{n-1} &= x^{n-1} + x^{n-2} w + \cdots + x w^{n-2} + w^{n-1} \\ &\quad - x^{n-1} - x^{n-1} - \cdots - x^{n-1} - x^{n-1} \\ &= \sum_{k=1}^{n-1} \left(x^{n-1-k} w^k - x^{n-1} \right) \\ &= \sum_{k=1}^{n-1} x^{n-1-k} \left(w^k - x^k \right). \end{aligned}$$

Let

$$h_n(w) = a_n \sum_{k=1}^{n-1} x^{n-1-k} \left(w^k - x^k \right) \quad \text{for } n = 2, 3, \cdots.$$

Then

$$\frac{f(w) - f(x)}{w - x} - g(x) = \sum_{n=2}^{\infty} h_n(w)$$

All h_n are continuous in $\mathbb{R}$ as they are polynomials in w; and $h_n(x) = 0$ for all $n \geqslant 2$. We claim that $\sum_{n=2}^{\infty} h_n(w)$ converges uniformly in $|w| \leqslant r$. In fact

$$\begin{aligned} |h_n(w)| &\leqslant |a_n| \sum_{k=1}^{n-1} |x|^{n-1-k} \left(|w|^k + |x|^k \right) \\ &\leqslant 2n |a_n| r^{n-1}. \end{aligned}$$

Now $\sum n|a_n|r^{n-1}$ is convergent, so that $\sum_{n=2}^{\infty} h_n(w)$ converges uniformly in closed disk $\{w : |w| \leqslant r\}$ by the Weierstrass M-test. Hence $\sum_{n=2}^{\infty} h_n(w)$ is continuous in the disk $|w| \leqslant r$ as the uniform limit of continuous functions is continuous. Therefore

$$\lim_{w \to x} \sum_{n=2}^{\infty} h_n(w) = \sum_{n=2}^{\infty} h_n(x) = 0$$

so that

$$\begin{aligned} \lim_{w \to x} \frac{f(w) - f(x)}{w - x} &= \lim_{w \to x} \left(\frac{f(w) - f(x)}{w - x} - g(x) \right) + g(x) \\ &= \lim_{w \to x} \sum_{n=2}^{\infty} h_n(w) + g(x) \\ &= g(x). \end{aligned}$$

□

Alternative Method of Proof: Alternatively, to prove the second part, we can apply Theorem 9.8 (for series). Let $f_n(x) = a_n x^n$, so that $f_n'(x) = n a_n x^{n-1}$ exists and is continuous. Further $\sum_{n=0}^{\infty} a_n x^n$ is convergent for $|x| < R$ and, for any $0 < r < R$, $\sum_{n=1}^{\infty} n a_n x^{n-1}$ is uniformly convergent on $[-r, r]$. Hence, by Theorem 9.8, for all $|x| \leq r$, $\sum_{n=0}^{\infty} a_n x^n$ is differentiable and

$$\frac{\mathrm{d}}{\mathrm{dx}} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

It follows that this is true for all $|x| < R$.

11.2 The Exponential Function, Trigonometric Functions, Hyperbolic Functions

The following result follows immediately

Proposition 11.3. *The functions* $\exp x$, $\sin x$, $\cos x$, $\cosh x$ *and* $\sinh x$ *can all be defined by power series with infinite radius of convergence so are all differentiable on* $\mathbb{R}$. *Further:*

(i) $\exp' x = \exp x$.

(ii) $\cos' x = -\sin x$ *and* $\sin' x = \cos x$.

(iii) $\cosh' x = \sinh x$ *and* $\sinh' x = \cosh x$.

Note 11.4. *The other trigonometric and hyperbolic functions are defined in terms of* cos *and* sin *or* cosh *and* sinh. *For example* $\tan x := \dfrac{\sin x}{\cos x}$ *is defined for those* x *such that* $\cos x \neq 0$. *Then by the quotient rule it is differentiable wherever it is defined, and* $\tan' x = \dfrac{\cos^2 x + \sin^2 x}{\cos^2 x}$. *We will soon give an easy proof that* $\cos^2 x + \sin^2 x = 1$[10]

11.3 Differentiability of the Inverse Function

Theorem 11.5 (Inverse Function Theorem (IFT)**).** *Let* $f : [a, b] \to [m, M]$ *be a strictly increasing continuous function from* $[a, b]$ *onto* $[m, M]$*, with inverse function* $g : [m, M] \to [a, b]$. *Suppose that* f *is differentiable at* $x_0 \in (a, b)$ *and that* $f'(x_0) \neq 0$. *Then* g *is differentiable at* $f(x_0)$*, and*

$$g'(f(x_0)) = \frac{1}{f'(x_0)}$$

Proof. We have already proved that g is continuous. Write $y_0 = f(x_0)$. Then for $y \neq y_0$

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}}$$

where $x = g(y)$, and so $y = f(x)$.

Since g is continuous, $x = g(y) \to g(y_0) = x_0$ as $y \to y_0$. Hence

$$\frac{f(x) - f(x_0)}{x - x_0} \to f'(x_0) \quad \text{as } \ y \to y_0.$$

As $f'(x_0) \neq 0$ we use AOL to see that

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \frac{1}{f'(x_0)}$$

exists. That is, g is differentiable at y_0, and

$$g'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

□

11.4 Logarithms

We continue to deal only with the real case where, in section 6, we defined $\log : (0, \infty) \to \mathbb{R}$ as the inverse function of the real exponential function.

[10]The Pythagoras Theorem.

To see that log is differentiable at any $y > 0$ proceed as we did when we discussed continuity, by finding an A such that $\exp(-A) < y < \exp(A)$ and then using the Inverse Function Theorem on the differentiable function $\exp : [-A, A] \to [\exp(-A), \exp(A)]$. We will find that

$$\log' y = \frac{1}{\exp'(\log y)} = \frac{1}{\exp(\log y)} = \frac{1}{y}$$

as we expect.

11.5 Powers

For any $x > 0$ and any $\alpha \in \mathbb{R}$ in section 6 we defined $x^\alpha = \exp(\alpha \log x)$. From the Chain Rule and the properties of exponentials and logarithms we therefore have that

$$\frac{\mathrm{d}}{\mathrm{d}x} x^\alpha = \alpha x^{\alpha - 1}.$$

12 Rolle's Theorem and the Mean Value Theorem

12.1 Local maxima and minima

Definition 12.1. *Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$.*

*(i) $x_0 \in E$ **is a local maximum** if for some $\delta > 0$, $f(x) \leqslant f(x_0)$ whenever $x \in (x_0 - \delta, x_0 + \delta) \cap E$.*

*(ii) $x_0 \in E$ **is a local minimum** if for some $\delta > 0$, $f(x) \geqslant f(x_0)$ whenever $x \in (x_0 - \delta, x_0 + \delta) \cap E$.*

A local maximum or minimum is called **a local extremum**. If the inequality is strict (for $x \neq x_0$) we will say that the extremum is strict.

Here is the crucial property (which, of course, you have met before).

Proposition 12.1 (Fermat's theorem on stationary points)**.** *Let $f : (a, b) \to \mathbb{R}$. Suppose that $x_0 \in (a, b)$ is a local extremum and f is differentiable at x_0. Then $f'(x_0) = 0$.*

Proof. If x_0 is a local maximum, then there exists $\delta > 0$ such that whenever $0 < x - x_0 < \delta$ and $x \in (a, b)$,

$$\frac{f(x) - f(x_0)}{x - x_0} \leqslant 0 \quad \text{so that}$$

$$f'_+(x_0) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0} \leqslant 0.$$

On the other hand, whenever $-\delta < x - x_0 < 0$ and $x \in (a, b)$,

$$\frac{f(x) - f(x_0)}{x - x_0} \geqslant 0 \quad \text{so that}$$

$$f'_-(x_0) = \lim_{x \to x_0 -} \frac{f(x) - f(x_0)}{x - x_0} \geqslant 0.$$

Since f is differentiable at x_0, $f'(x_0) = f'_-(x_0) = f'_+(x_0)$ and hence $f'(x_0) = 0$.

Similarly if x_0 is a local minimum. □

Remark 12.2. *It is essential that the interval (a, b) is open. Why?*

12.2 Rolle's Theorem

Theorem 12.3 (Rolle, 1691)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, and differentiable on (a, b). Suppose further that $f(a) = f(b)$. Then there exists a point $\xi \in (a, b)$ such that $f'(\xi) = 0$.*

Proof. If f is constant in $[a, b]$, then $f'(x) = 0$ for every $x \in (a, b)$, so that any point—say $\xi = \frac{1}{2}(a + b)$—will do.

As f is continuous on $[a, b]$ it attains its maximum and minimum on $[a, b]$ (by Theorems 4.1 and 4.4). As $f(a) = f(b)$, either f is constant and we are done, or else the maximum or the minimum lies in the open interval (a, b). Suppose that $\xi \in (a, b)$ gives either the maximum or minimum. Then it is a local extremum, and by Fermat's result $f'(\xi) = 0$. □

We can express this informally by saying

'between any two roots of f there is a root of f''.

Note 12.4. *(i) Remember that f is differentiable implies that f is continuous. Thus the hypotheses of Rolle would be satisfied if f was differentiable on $[a, b]$ and $f(a) = f(b)$. However, often it is important that Rolle holds under the given weaker conditions.*

(ii) When using these theorems remember to check ALL conditions including the continuity and differentiability conditions. For example $f : [-1, 1] \to \mathbb{R}$ given by $f(x) = |x|$ satisfies all conditions of Rolle except that f is not differentiable at $x = 0$. But there is no ξ such that $f'(\xi) = 0$.

12.3 The Mean Value Theorem

This is one of the most important results in this course. It is a rotated version of Rolle.[11]

[11] If in an examination you are asked to prove the Mean Value Theorem, then you need to provide also proofs of Fermat's result and Rolle's Theorem.

Theorem 12.5 (MVT)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, and differentiable on (a, b). Then there exists a point $\xi \in (a, b)$ such that*

$$f(b) - f(a) = f'(\xi)(b - a).$$

Proof. Apply Rolle's theorem to the function

$$F(x) = f(x) - k(x - a),$$

where k is a constant to be determined. $F : [a, b] \to \mathbb{R}$ is continuous, and is differentiable on (a, b). We choose k so that $F(a) = F(b)$, that is $k = \dfrac{f(b) - f(a)}{b - a}$. Thus Rolle's theorem applies, so for some number $\xi \in (a, b)$, $F'(\xi) = 0$. But $F'(x) = f'(x) - k$, so $f'(\xi) = k = \dfrac{f(b) - f(a)}{b - a}$, as required. □

Note 12.6. *Suppose we have the hypotheses of the MVT. Then for any $a \leqslant a_1 < b_1 \leqslant b$ we can apply the MVT to f restricted to $[a_1, b_1]$ and get*

$$f(b_1) - f(a_1) = f'(\xi_1)(b_1 - a_1) \qquad \text{for some } \xi_1 \in (a_1, b_1).$$

Note that (for a given function f) the value of ξ_1 may depend on a_1 and b_1.

Corollary 12.7 (Taylor's Theorem, mark 1)**.** *Suppose that we have the hypotheses of the MVT and that $x, x + h \in [a, b]$. Then*

$$f(x + h) - f(x) = f'(x + \theta h)h \qquad \text{for some } \theta \in (0, 1).$$

Proof. Suppose $h < 0$; then $a \leqslant x + h < x \leqslant b$. From the MVT applied to f on the interval $[x + h, x]$ there exists $\xi \in (x + h, x)$ such that

$$f(x) - f(x + h) = f'(\xi)(-h).$$

Write $\xi = x + \theta h$, and note that $x + h < x + \theta h < x$ implies—as $h < 0$—that $0 < \theta < 1$.

The cases $h = 0$ and $h > 0$ are left as exercises. □

12.4 A Function with Zero Derivative is Constant

Here is one of the most useful consequences of the MVT.

Corollary 12.8 (Constancy Theorem - A function with zero derivative is constant)**.** *Let $f : (a, b) \to \mathbb{R}$ be differentiable, and satisfy $f'(t) = 0$ for all $t \in (a, b)$. Then f is constant on (a, b).*

Proof. Apply MVT to f on $[x, y]$ where x, y are any two points in (a, b). (Note that f is differentiable on (a, b) implies that f is continuous on (a, b) and hence f is continuous on $[x, y]$.) Then $f(x) - f(y) = f'(\xi)(x - y)$ for some $\xi \in (x, y)$. But $f'(\xi) = 0$, so that $f(x) = f(y)$. Therefore f is constant in (a, b). □

Note that the interval (a, b) need not be bounded.

Example 12.1. *Suppose that ϕ is a function whose derivative is x^2. Then we have, for all x, that $\phi(x) = \frac{1}{3}x^3 + A$ for some constant A.*

Proof. Let $f(x) := \phi(x) - \frac{1}{3}x^3$; then f is differentiable and we can calculate that $f'(x) = x^2 - \frac{1}{3} \cdot 3x^2 = 0$. By the Constancy Theorem $f(x) = A$ for some constant A. You can justify other 'integrations' similarly . Just guess the 'integral' and proceed as above. □

Note 12.9. *Solutions of differential equations: Last term you learned methods for guessing solutions of first and second order linear odes. You were then told that these solutions could be used to get the general solution. The Constancy Theorem gives us a tool to prove the uniqueness of solutions of DEs and to justify that you did indeed have general solutions last term as was claimed (see also Section 13.3). Those who do PDEs have already seen this idea this term where you showed that $E'(t) = 0$ and then deduced that $E(t)$ is a constant (which then turned out to be zero).*

Here is a very fundamental example of how we use the MVT to find the general solution of a differential equation.

Example 12.2. *Show that the general solution for $f'(x) = f(x)$ for all $x \in \mathbb{R}$, is $f(x) = A\exp(x)$ where A is a constant. (i.e. every solution is of this form)*

The 'trick' for solving differential equations is to manipulate them so that they look like $\dfrac{\mathrm{d}}{\mathrm{d}x}F = 0$ for some F, and then 'integrate'. This can often be achieved by multiplying by 'integrating factors'. The same 'trick' lets us apply the MVT (or the Constancy Theorem) to *prove* that the solution must be of this form.

Last term you learnt that to solve the differential equation $\dfrac{\mathrm{d}f}{\mathrm{d}x} - f = 0$ you multiply it by $\mathrm{e}^{-\int 1\mathrm{d}x}$, rewrite it as $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-x}f(x)\right) = 0$ and deduce that $\mathrm{e}^{-x}f(x) = A$.

Now write this as a piece of pure mathematics!

Consider $F(x) := f(x)\exp(-x)$. Then $F'(x) = f'(x)\exp(-x) - f(x)\exp(-x) = 0$. Hence, by the Constancy Theorem $F(x)$ is constant; that is $f(x)\exp(-x) = A$ say, and so $f(x) = A\exp(x)$ and all solutions are of this form.

12.5 Derivatives and monotonicity

Corollary 12.10. *Let $f : (a, b) \to \mathbb{R}$ be differentiable.*

(i) If $f'(x) \geqslant 0$ for all $x \in (a, b)$ then f is increasing on (a, b).

Proof: Apply the MVT to any $[x, y] \subset (a, b)$ to get $f(y) - f(x) = f'(\xi)(y - x)$, a product of non-negative numbers. Hence $f(y) \geqslant f(x)$ and we are done.

(ii) If $f'(x) \leqslant 0$ for all $x \in (a, b)$ then f is decreasing on (a, b).

(iii) If $f'(x) > 0$ for all $x \in (a, b)$ then f is strictly increasing on (a, b).

(iv) If $f'(x) < 0$ for all $x \in (a, b)$ then f is strictly decreasing on (a, b).

12.6 The Cauchy Mean Value Theorem

Sometimes we are concerned with more than one function, and would like to use the MVT or a MVT type argument. The following is what we need: except in the most trivial cases it never helps to apply the MVT to the functions separately—we generate too many distinct ξ's.

Corollary 12.11 (Cauchy's Mean Value Theorem)**.** [12] *Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Suppose that $g'(x) \neq 0$ for all $x \in (a, b)$. Then for some $\xi \in (a, b)$ we have that*

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof. To be supplied later when we need the result. □

13 Applications of the MVT

Now we have the MVT we can deduce many properties of the exp, log and trig functions very easily.

13.1 Exponential and Logarithm

Proposition 13.1. $\exp(x + y) = \exp(x)\exp(y)$ *for all* $x, y \in \mathbb{R}$.

Proof. We will use the Constancy Theorem—but on what function? Fixing y and looking at $f(x) = \exp(x + y) - \exp(x)\exp(y)$ leads to $f' = f$ and $f(0) = 0$ which we could now solve to get $f(x) = 0$ (see section 12.4).

[12] This is where the result belongs logically, but in the lectures it will not appear until later, when we do L'Hospital's Rule.

However a much better (more direct) way is to fix $x + y$ instead. So, fix $c \in \mathbb{R}$, and put $g(t) = \exp c - \exp t \exp(c - t)$. Then we have that $g'(t) = 0$ so that $g(t) = g(0)$ by the Constancy Theorem. Now $g(0) = \exp c - \exp 0 \exp c = 0$. So for any c, t we have that $\exp c - \exp t \exp(c - t) = 0$. Put $c := (x + y)$, and $t := x$ to get the result. □

Corollary 13.2. $\log(uv) = \log(u) + \log(v)$ *for all* $u, v \in (0, \infty)$.

Proof. From above

$$\exp(\log(u) + \log(v)) = \exp(\log(u)) \exp(\log(v)) = uv = \exp(\log(uv))$$

and take logs. □

We can also use the MVT to prove the monotonicity of the exponential function.

Proposition 13.3. *The function* $\exp : \mathbb{R} \to (0, \infty)$ *is strictly increasing.*

Proof. As $\exp x > 0$, its derivative is positive. □

13.2 Trigonometric Functions

Proposition 13.4 (The Pythagoras Theorem)**.** *For all real* x *we have that*

$$\cos^2 x + \sin^2 x = 1.$$

Proof. Let $f(x) := \cos^2 x + \sin^2 x - 1$. Then by what we have proved about derivatives of trigonometric functions,

$$f'(x) = 2 \cos x(-\sin x) + 2 \sin x \cos x - 0 = 0.$$

for all x.

By the Constancy Theorem

$$f(x) = f(0) = \cos^2 0 + \sin^2 0 - 1 = (1)^2 - 0 - 1 = 0.$$

as required. □

Proposition 13.5 (Addition Formulae)**.** *For all real* x, y *we have that*

(i) $\cos(x + y) = \cos x \cos y - \sin x \sin y$*;*

(ii) $\sin(x + y) = \sin x \cos y + \cos x \sin y$.

Proof. It is enough to prove one, the other is got by fixing y and taking the derivative of the resulting function of x.

To prove (i) we recall what we did for exponentials: let

$$h(x) = \cos c - \cos x \cos(c - x) + \sin x \sin(c - x)$$

whose derivative is

$$h'(x) = 0 + \sin x \cos(c - x) - \cos x \sin(c - x) + \cos x \sin(c - x) - \sin x \cos(c - x) = 0$$

so that by the Constancy Theorem $h(x) = h(c) = 0$. $\square$

Proposition 13.6. *The function* $\cos x := \sum_0^\infty \dfrac{(-1)^k x^{2k}}{(2k)!}$ *has a least positive zero which we denote (for the moment) by* α.

Proof. First we need to see that there are positive zeros. Note that

$$\cos 0 = \sum_0^\infty \frac{(-1)^k 0^{2k}}{(2k)!} = 1 > 0$$

and (by looking at pairs of terms)

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \sum_1^\infty \frac{x^{4k+2}}{(4k+2)!}\left(1 - \frac{x^2}{(4k+4)(4k+3)}\right) \leqslant 1 - \frac{x^2}{2!} + \frac{x^4}{4!}$$

provided $x^2 \leqslant (4{+}4)(4{+}3)$. As $1 - \frac{x^2}{2!} + \frac{x^4}{4!} = \frac{1}{24}\left[(x^2-6)^2 - 12\right]$ we see that $\cos\sqrt{6} < 0$. By the IVT, $\cos x$ has at least one zero in $[0, \sqrt{6}]$.

Now let

$$S = \{t > 0 : \cos t = 0\}.$$

Then $S \neq \emptyset$ and S is bounded below, so that $\alpha = \inf S$ exists. By definition of $\inf S$, given n there exists $t_n \in S$ such that $\alpha \leq t_n < \alpha + 1/n$. Thus $t_n \to \alpha$ as $n \to \infty$. But $\cos x$ is continuous, so that $\cos t_n \to \cos\alpha$, and hence $\cos\alpha = 0$. But $\cos 0 = 1$ so that α is the minimum positive zero required. $\square$

Proposition 13.7. $\sin\alpha = 1$.

Proof. By Pythagoras, $\sin\alpha = \pm 1$. Suppose $\sin\alpha = -1$; then by the MVT there would be some $\xi \in (0, \alpha)$ such that $\cos\xi = \dfrac{\sin\alpha - \sin 0}{\alpha - 0} = \dfrac{-1}{\alpha} < 0$. However, $\cos 0 = 1$, and α is the first root, so by the IVT $\cos\xi$ cannot be negative. $\square$

Proposition 13.8 (Periodicity). *For all real* x *we have that*

(i) $\cos(x + \alpha) = -\sin x$ *and* $\sin(x + \alpha) = \cos x$;

(ii) $\cos(x + 2\alpha) = -\cos x$ *and* $\sin(x + 2\alpha) = -\sin x$*;*

(iii) $\cos(x + 4\alpha) = \cos x$ *and* $\sin(x + 4\alpha) = \sin x$*.*

(iv) $\cos(2k\alpha) = (-1)^k$*;* $\cos((2k + 1)\alpha) = 0$*;*
$\sin(2k\alpha) = 0$*;* $\sin((2k + 1)\alpha) = (-1)^k$*,* $\forall k \in \mathbb{Z}$*.*

Proof. We just use the addition formula repeatedly, inserting the values $\cos\alpha = 0$ and $\sin\alpha = 1$. $\square$

Now that we have proved these results, and the danger of using 'obvious' but unproved properties of π has passed we can make the following definition:

Definition 13.1. $\pi := 2 \cdot \inf\{t > 0 : \cos t = 0\}(= 2\alpha)$.

We need one more result, and then we have established "all" the usual facts about the trigonometric functions.

Proposition 13.9. *The zeros of* $\cos x$ *are at precisely the points* $\{(k + \frac{1}{2})\pi \ : \ k \in \mathbb{Z}\}$*.*

Proof. By Proposition 13.8(iv), for $k \in \mathbb{Z}$, $\cos(\frac{1}{2}\pi + k\pi) = 0$ so these are all zeros. If β is such that $\cos\beta = 0$ then from Proposition 13.8(ii)there exists $k \in \mathbb{Z}$ such that $\beta_0 = \beta + k\pi \in (0, \pi]$ is a zero of $\cos x$. Clearly $\beta_0 \not< \frac{1}{2}\pi$ by definition. Using

$$\cos(\pi - x) = -\cos(-x) = -\cos(x)$$

we see that if $\beta_0 > \frac{1}{2}\pi$ then $\pi - \beta_0 < \frac{1}{2}\pi$ is a zero of $\cos x$, which cannot be. Hence $\beta_0 = \frac{1}{2}\pi$, and β has the required form. $\square$

Note that from Proposition 13.8(i) it now follows that the zeros of $\sin x$ are precisely $\{k\pi \ : \ k \in \mathbb{Z}\}$.

13.3 Differential Equations

In Section 12 we already looked at one example. Here is another, but this will not be covered in lectures.

Example 13.1. *This example is based on a Calculus question from a Mods Collection. Find all solutions of*

$$y'' - \frac{2}{1 + x^2} y = 0.$$

(The emphasis for us is on "all".)

The following is all motivated by the method for finding a second solution for second-order linear ordinary differential equations when one solution is known, which you learnt in the 'Introductory Calculus' course last term. We use the methods from this course to show that these are **all** the solutions.

We can check easily that $(1+x^2)$ is a solution; so we write

$$z(x) = \frac{y(x)}{1+x^2}.$$

An easy calculation yields

$$y' = z'(1+x^2) + 2xz \text{ and } y'' = z''(1+x^2) + 4xz' + 2z,$$

so that z must satisfy

$$z''(1+x^2) + 4xz' = 0$$

and hence

$$\left[z'(1+x^2)^2\right]' = (z''(1+x^2) + 4xz')(1+x^2) = 0.$$

By the Constancy Theorem.

$$z'(1+x^2)^2 = A$$

for some constant A and so

$$z'(x) = \frac{A}{(1+x^2)^2}.$$

Although of course we can't 'integrate up' yet — we don't know what that means — we can take the hint and look at what the integral would be, namely

$$w(x) = \frac{1}{2}\left[\arctan x + \frac{x}{1+x^2}\right];$$

here arctan is the inverse function of tan. So by the Inverse Function Theorem and the other rules of differentiation which we have established we can check that

$$w'(x) = \frac{1}{(1+x^2)^2} = z'(x)/A.$$

Hence by the Constancy Theorem $z(x) - Aw(x) = B$ for some constant B, and so the **only** solutions are

$$y(x) = \frac{A}{2}\left[(1+x^2)\arctan x + x\right] + B(1+x^2).$$

13.4 The function $\dfrac{\sin x}{x}$

This is a good example of how the Mean Value Theorem and its various corollaries are used practically.

Proposition 13.10. *Let* $0 < x < \frac{1}{2}\pi$. *Then*

(i) $\sin x < x < \tan x$ *and so* $\cos x < \frac{\sin x}{x} < 1$*;*

(ii) $\lim_{x\to 0} \frac{\sin x}{x} = 1$*;*

(iii) $\frac{2}{\pi} < \frac{\sin x}{x} < 1$ *(Jordan's inequality).*

We therefore have the following bounds:

$$\max\{\cos x, \frac{2}{\pi}\} < \frac{\sin x}{x} < 1.$$

Proof. To prove the first inequality, consider $f(x) = \tan x - x$, for $x \in [0, \frac{1}{2}\pi)$. Then f is differentiable on $(0, \frac{1}{2}\pi)$ and

$$f'(x) = \frac{1}{\cos^2 x} - 1 > 0 \quad \text{for all } x \in (0, \tfrac{1}{2}\pi).$$

Hence f is strictly increasing on $[0, \frac{1}{2}\pi)$; in particular $f(x) > f(0)$ for any $x \in (0, \frac{1}{2}\pi)$ which yields $\tan x > x$. Considering $x - \sin x$ in the same way will give $x > \sin x$.

The second inequality in (i) is got by inverting and multiplying by $\sin x$; this is justified since $\sin x > 0$ until the smallest positive zero of $\cos x$.

For (ii) we use a version of the sandwich theorem and the continuity of $\cos x$ to get that $\lim_{x\to 0+} \frac{\sin x}{x}$ exists and

$$1 = \lim_{x\to 0+} \cos x \leqslant \lim_{x\to 0+} \frac{\sin x}{x} \leqslant 1.$$

As $\dfrac{\sin x}{x}$ is an even function this gives that $\lim_{x\to 0} \dfrac{\sin x}{x} = 1$.

Now consider

$$h(x) = \frac{\sin x}{x} \quad \text{for } x \in (0, \tfrac{1}{2}\pi].$$

Then

$$h'(x) = \frac{\cos x(x - \tan x)}{x^2} < 0 \quad \text{for all } x \in (0, \tfrac{1}{2}\pi)$$

so that h is strictly decreasing, and hence $h(x) > h(\frac{1}{2}\pi)$ for any $x \in (0, \frac{1}{2}\pi)$; this gives the first inequality of (iii). The second is already included in (i).

□

14 L'Hôpital's Rule

This section is devoted to a variety of rules and techniques for calculating limits of quotients. They derive from results of Guillaume de l'Hôpital; perhaps they are really due to Johann Bernoulli whose lecture notes l'Hôpital published in 1696.

14.1 The Cauchy Mean Value Theorem

As promised earlier here is the proof of Cauchy's symmetric form of the MVT. (At first sight one might think we could just apply the MVT to f and g separately. However, a moment's reflection will show that we would then get two different ξ.)

Theorem 14.1 (Cauchy's Mean Value Theorem)**.** *Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Suppose that $g'(x) \neq 0$ for all $x \in (a, b)$. Then for some $\xi \in (a, b)$ we have that*

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof. First, this makes sense: we cannot have $g(b) - g(a) = 0$ or by Rolle's Theorem there would be a point $\eta \in (a, b)$ with $g'(\eta) = 0$.

Now let the function F be defined on $[a, b]$ by

$$F(x) := \begin{vmatrix} 1 & 1 & 1 \\ f(x) & f(a) & f(b) \\ g(x) & g(a) & g(b) \end{vmatrix}$$

that is

$$F(x) = (f(a)g(b) - f(b)g(a)) + f(x)\,(g(a) - g(b)) + g(x)\,(f(b) - f(a))$$

which, being a linear combination of f and g is continuous on $[a, b]$ and differentiable on (a, b). Clearly $F(a) = F(b) = 0$; so Rolle's Theorem applies and yields a $\xi \in (a, b)$ such that $F'(\xi) = 0$. But

$$0 = F'(\xi) = 0 + f'(\xi)\,(g(a) - g(b)) + g'(\xi)\,(f(b) - f(a))$$

and we are done after dividing by the non-zero $g'(\xi)(g(b) - g(a))$. $\square$

14.2 The L'Hôpital Rule

Proposition 14.2. *Suppose f, g are continuous on $[a, a + \delta]$ (for some $\delta > 0$), and differentiable in $(a, a + \delta)$, and that $f(a) = g(a) = 0$. Suppose further that $l := \lim_{x \to a+} \frac{f'(x)}{g'(x)}$ exists.*

Then

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = \lim_{x \to a+} \frac{f'(x)}{g'(x)}.$$

Proof. Note that there must exist a $\delta' < \delta$ such that on $(a, a + \delta']$ we have that $g'(x) \neq 0$, for otherwise the function $f'(x)/g'(x)$ would not be defined near a and so this limit could not be defined.

For every $x \in (a, a+\delta')$, apply Cauchy's MVT to f, g on the interval $[a, x]$: there is $\xi_x \in (a, x)$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi_x)}{g'(\xi_x)}.$$

But if $x \to a+$, then $\xi_x \to a$ with $\xi_x > a$, so that

$$\lim_{x \to a+} \frac{f'(\xi_x)}{g'(\xi_x)} = l.$$

Hence

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = \lim_{x \to a+} \frac{f'(\xi_x)}{g'(\xi_x)} = l$$

□

Similarly we prove

Corollary 14.3. *Suppose f, g are continuous on $[a-\delta, a]$ (for some $\delta > 0$), and differentiable in $(a-\delta, a)$, and that $f(a) = g(a) = 0$. Suppose further that $l := \lim_{x \to a-} \frac{f'(x)}{g'(x)}$ exists. Then*

$$\lim_{x \to a-} \frac{f(x)}{g(x)} = \lim_{x \to a-} \frac{f'(x)}{g'(x)}.$$

The proof of the following is now immediate.

Corollary 14.4 (L'Hôpital's Rule (L'HR))**.** *Suppose f, g are continuous on $[a-\delta, a+\delta]$ (for some $\delta > 0$), and differentiable in $(a-\delta, a+\delta) \setminus \{a\}$, and that $f(a) = g(a) = 0$. Suppose further that $l := \lim_{x \to a} \frac{f'(x)}{g'(x)}$ exists. Then*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

Note 14.5. *Sometimes this is called the $\frac{0}{0}$ case of L'HR.*

14.3 Some Applications

Example 14.1. *Prove that*

$$\lim_{x \to 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}.$$

We argue like this:

$$
\begin{aligned}
\lim_{x\to 0}\frac{1-\cos x}{x^2} &= \lim_{x\to 0}\frac{\sin x}{2x} \quad \text{by L'HR, provided this limit exists}\\
&= \lim_{x\to 0}\frac{\cos x}{2} \quad \text{by L'HR, provided this limit exists}\\
&= \frac{1}{2} \quad \text{and this limit exists by the continuity of } \cos x;\\
&\qquad \text{so the above equalities hold.}
\end{aligned}
$$

To justify all this we need to check L'HR, which we have used twice is actually applicable. But by standard results we have already proved:
$1-\cos x$ and x^2 are continuous on $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$, zero at zero, and differentiable on $(\frac{1}{2}\pi, \frac{1}{2}\pi)\setminus\{0\}$ with derivatives $\sin x$ and $2x$;
$\sin x$, and $2x$ are continuous on $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$, zero at zero, and differentiable on $(\frac{1}{2}\pi, \frac{1}{2}\pi)\setminus\{0\}$ with derivatives $\cos x$ and 2

Note that this proves incidentally that $\lim_{x\to 0}\frac{\sin x}{x} = 1$.

Example 14.2.

$$\lim_{x\to 0}\frac{\log(1+x)}{x} = 1.$$

Again we argue:

$$
\begin{aligned}
\lim_{x\to 0}\frac{\log(1+x)}{x} &= \lim_{x\to 0}\frac{\log'(1+x)}{x'} \quad \text{by L'HR, provided this limit exists}\\
&= \lim_{x\to 0}\frac{1}{1+x} \quad \text{derivative of } \log t \text{ is } \frac{1}{t}\\
&= 1 \quad \text{by continuity of } \frac{1}{1+x}\\
&\qquad \text{—as this exists previous equalities hold.}
\end{aligned}
$$

To justify the use of L'HR we need to see that $\log(1+x)$ and x are continuous on $[-\frac{1}{2},\frac{1}{2}]$, 0 at 0, and differentiable on $(-\frac{1}{2},\frac{1}{2})\setminus\{0\}$.

Example 14.3.

$$\lim_{x\to 0}(1+x)^{\frac{1}{x}} = \mathrm{e}.$$

Recall that by definition $(1+x)^{\frac{1}{x}} := \exp\left(\frac{1}{x}\log(1+x)\right)$. So consider first $\frac{\log(1+x)}{x}$. By the previous example this has limit 1. Now by the continuity of $\exp(x)$ we see that

$$(1+x)^{\frac{1}{x}} = \exp\left(\frac{\log(1+x)}{x}\right) \to \exp(1) = \mathrm{e} \quad \text{as } x\to 0.$$

14.4 L'Hôpital's Rule: infinite limits

If we have all the hypotheses for L'Hôpital's rule, except that we have

$$\frac{f'(x)}{g'(x)} \to +\infty \qquad \text{as} \;\; x \to a$$

then we swap f and g, then use L'HR and conclude that

$$\frac{g(x)}{f(x)} \to 0 \qquad \text{as} \;\; x \to a.$$

14.5 L'Hôpital's Rule at ∞

Proposition 14.6. *Suppose $f, g : (a, +\infty) \to \mathbb{R}$ are continuous and differentiable, with $f(x) \to 0$ and $g(x) \to 0$ as $x \to \infty$. If $g'(x) \neq 0$ on $(a, +\infty)$ and $\frac{f'(x)}{g'(x)} \to l$ as $x \to \infty$, then $\lim_{x\to\infty} \frac{f(x)}{g(x)} = l$.*

Sketch proof: Put $y = \frac{1}{x}$ so $y \to 0$ as $x \to \infty$. Then apply L'HR to the functions $F(y) = f(\frac{1}{y})$ and $G(y) = g(\frac{1}{y})$, with $F(0) = 0 = G(0)$, checking carefully that the hypotheses hold.

14.6 L'Hôpital's Rule—the $\frac{\infty}{\infty}$ case

There is one important variant which we cannot obtain by algebraic manipulation, or by taking logarithms or exponentials or similar tricks. The proof will probably not be covered in the lectures.

Proposition 14.7 (L'HR, the $\frac{\infty}{\infty}$ case)**.** *Let $f, g : (a, a+\delta) \to \mathbb{R}$ be differentiable for some $\delta > 0$. Suppose further that $f(x) \to \infty$ and $g(x) \to \infty$ as $x \to a+$ and that $\lim_{x\to a+} \frac{f'(x)}{g'(x)}$ exists.*

Then

$$\lim_{x\to a+} \frac{f(x)}{g(x)} = \lim_{x\to a+} \frac{f'(x)}{g'(x)}.$$

Note 14.8. *We do not want to make too much heavy weather in this proof; checking all the details is a good exercise.*

Proof. Write $K := \lim_{x\to a+} \frac{f'(x)}{g'(x)}$. Let $\varepsilon > 0$, then there exists a $\delta_1 > 0$ such that $\delta_1 < \delta$ and

$$\left| \frac{f'(x)}{g'(x)} - K \right| < \tfrac{1}{2}\varepsilon \quad \text{for all } x \in (a, a+\delta_1).$$

Now fix some c in $(a, a+\delta_1)$.

For any $x \in (a, c)$ we apply Cauchy's MVT to f, g on $[x, c]$: there is a number $\xi_x \in (x, c)$ such that

$$\frac{f(c) - f(x)}{g(c) - g(x)} = \frac{f'(\xi_x)}{g'(\xi_x)}.$$

Since $\xi_x \in (x, c) \subset (a, a + \delta_1)$, we have that

$$\left| \frac{f(x) - f(c)}{g(x) - g(c)} - K \right| = \left| \frac{f'(\xi_x)}{g'(\xi_x)} - K \right| < \tfrac{1}{2}\varepsilon \quad \text{for all } x \in (a, c).$$

(Unlike the $\frac{0}{0}$ case we cannot conclude immediately that $\frac{f(x)-f(c)}{g(x)-g(c)} \to K$ as $x \to a+$ (although it does !!), as there is no guarantee that ξ_x will tend to a as $x \to a+$).

Clearing the fraction we have that

$$|f(x) - f(c) - Kg(x) + Kg(c)| < \tfrac{1}{2}\varepsilon\, |g(x) - g(c)|$$

so that the Triangle Law gives us

$$|f(x) - Kg(x)| < \tfrac{1}{2}\varepsilon\, |g(x) - g(c)| + |f(c) - Kg(c)|$$

or, provided $g(x) \neq 0$

$$\left| \frac{f(x)}{g(x)} - K \right| < \tfrac{1}{2}\varepsilon \left| 1 - \frac{g(c)}{g(x)} \right| + \frac{|f(c) - Kg(c)|}{|g(x)|}.$$

Now use the fact that $g(x) \to \infty$; we can find a $\delta_2 > 0$, such that $\delta_2 < \delta_1$ and such that for $a < x < a + \delta_2$, $g(x) \neq 0$,

$$\left| 1 - \frac{g(c)}{g(x)} \right| < \tfrac{3}{2} \quad \text{and} \quad \frac{|f(c) - Kg(c)|}{|g(x)|} < \tfrac{1}{4}\varepsilon$$

so that we have

$$\left| \frac{f(x)}{g(x)} - K \right| < \tfrac{1}{2}\tfrac{3}{2}\varepsilon + \tfrac{1}{4}\varepsilon = \varepsilon$$

as required. $\square$

14.7 More applications

These examples might be better done by using the standard limit from Analysis I that, if $\alpha > 0$, $x \exp(-\alpha x) \to 0$ as $x \to \infty$.

Example 14.4. $\lim_{x \to +\infty} \frac{\log x}{x^\mu} = 0$ *for any* $\mu > 0$.

Let $g(x) = x^\mu = \exp(\mu \log x)$. Then $g'(x) = \mu x^{\mu-1}$. So by L'Hôpital's rule ($\frac{\infty}{\infty}$ case) we have

$$\begin{aligned} \lim_{x \to +\infty} \frac{\log x}{x^\mu} &= \lim_{x \to +\infty} \frac{\frac{1}{x}}{\mu x^{\mu-1}} \quad \text{provided this limit exists} \\ &= \lim_{x \to +\infty} \frac{1}{\mu x^\mu} = 0 \quad \text{which } \textit{does} \text{ exist.} \end{aligned}$$

Example 14.5. *For any* $\mu > 0$, $\lim_{x\to 0+} x^\mu \log x = 0$.

We transform this into $\frac{\infty}{\infty}$ form and then by L'HR

$$\begin{aligned}
\lim_{x\to 0+} x^\mu \log x &= \lim_{x\to 0+} \frac{\log x}{x^{-\mu}} \\
&= \lim_{x\to 0+} \frac{\log' x}{(x^{-\mu})'} \quad \text{if this limit exists} \\
&= \lim_{x\to 0+} \frac{\frac{1}{x}}{(-\mu)x^{-\mu-1}} = \lim_{x\to 0+} \frac{x^\mu}{(-\mu)} = 0 \quad \text{which } \textit{does} \text{ exist.}
\end{aligned}$$

Finally

Example 14.6. *Show that*

$$\lim_{x\to 0} \left(\tfrac{\sin x}{x}\right)^{\frac{1}{1-\cos x}} = e^{-\frac{1}{3}}.$$

Since $f(x) = \left(\frac{\sin x}{x}\right)^{\frac{1}{1-\cos x}}$ is an even function, we only need to show that $\lim_{x\to 0+} f(x) = e^{-\frac{1}{3}}$. According to the definition

$$\begin{aligned}
f(x) &= \exp\left(\frac{1}{1-\cos x} \log \frac{\sin x}{x}\right) \\
&= \exp\left(\frac{\log \sin x - \log x}{1-\cos x}\right) .
\end{aligned}$$

By the L'Hôpital Rule,

$$\begin{aligned}
\lim_{x\to 0+0} \frac{\log \sin x - \log x}{1-\cos x} &= \lim_{x\to 0+} \frac{\frac{\cos x}{\sin x} - \frac{1}{x}}{\sin x} \quad \text{[provided it exists; recall} \frac{\sin x}{x} \to 1] \\
&= \lim_{x\to 0+} \frac{x\cos x - \sin x}{x \sin^2 x} \\
&= \lim_{x\to 0+} \frac{\cos x - x\sin x - \cos x}{\sin^2 x + 2x \sin x \cos x} \text{[if it exists,using L'Hôpital]} \\
&= -\lim_{x\to 0+} \frac{x}{\sin x + 2x\cos x} \\
&= -\lim_{x\to 0+} \frac{1}{\cos x + 2\cos x - 2x \sin x} \text{[if it exists,using L'Hôpital]} \\
&= -\frac{1}{3} \quad \text{[continuity]} .
\end{aligned}$$

Finally, since exp is continuous at $-\frac{1}{3}$,

$$\begin{aligned}
\lim_{x\to 0+} \left(\frac{\sin x}{x}\right)^{\frac{1}{1-\cos x}} &= \lim_{x\to 0+} \exp\left(\frac{\log \sin x - \log x}{1-\cos x}\right) \\
&= \exp\left(-\frac{1}{3}\right) \quad \text{[by continuity of exp]} .
\end{aligned}$$

14.8 Health Warning

L'Hôpital's Rule is very seductive. But it is often *not* the best way to evaluate limits. Taylor's Theorem, to which we turn next, is often more useful, and indeed more informative.

If you doubt this, then use L'HR to work out $\lim_{x\to 0} \dfrac{\sinh x^4 - x^4}{(x - \sin x)^4}$, and then later use Taylor's Theorem to write it down at sight—and decide which is better.

15 Taylor's Theorem

15.1 Motivation

Suppose that $f : (a-\delta, a+\delta) \to \mathbb{R}$ and that for some $n \geqslant 1$ the derivatives f', f'', $\ldots$, $f^{(n)}$ exist on the interval. For convenience write $f^0 := f$.

We can then form the *Taylor polynomials*

$$P_n(x) := f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

a polynomial of degree n in x. This polynomial 'agrees with f' to the extent that $P_n^{(k)}(a) = f^{(k)}(a)$ for $k = 0, \ldots, n$.

We have

$$\begin{aligned}
P_0(x) &= f(a) \quad \text{constant approximation, not very interesting;}\\
P_1(x) &= f(a) + f'(a)(x-a) \quad \text{linear approximation;}\\
P_2(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 \quad \text{quadratic approximation;}\\
&\cdots \quad \text{and so on.}
\end{aligned}$$

We might hope, on the basis of our experience, that $P_n(x)$ is a good approximation to $f(x)$; we would like to investigate that intuition.

We will also consider the power series

$$P(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

which is called the Taylor expansion of f at a. Our previous experience leads us to conjecture that this must equal $f(x)$.

To investigate these questions we will look at the 'error term'

$$E_n(x) := f(x) - P_n(x).$$

(Clearly, if f has derivatives of all orders, $P_n(x) \to f(x)$ as $n \to \infty$ if and only if $E_n(x) \to 0$.) Unfortunately, even if f has derivatives of all orders, it need not be true that $E_n(x) \to 0$ as $n \to \infty$, so we have to move more carefully. First, we will prove Taylor's Theorem which will give us information about $E_n(x)$. Secondly, in individual cases we have to consider whether $E_n(x) \to 0$ as $n \to \infty$.

15.2 A cautionary example

Our intuition, built on experience of polynomials, trigonometric and exponential functions, is misleading. The following example shows us that there are functions $f(x)$, with derivatives of all orders at every point of $\mathbb{R}$, such that $\sum \frac{f^k(0)}{k!}x^k$ is convergent for every x—but for which $E_n(x) \not\to 0$.

Consider $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \exp(-\frac{1}{x^2}) & \text{whenever } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

Some experimentation shows that we expect

$$f^{(k)}(x) = \begin{cases} Q_k(\frac{1}{x})\exp(-\frac{1}{x^2}) & \text{whenever } x \neq 0, \\ 0 & \text{for } x = 0, \end{cases}$$

for some polynomial Q_k of degree $3k$. We can prove this by induction: At points $x \neq 0$ this is routine use of linearity, the product rule and the chain rule. But at $x = 0$ we need to take more care, and use the definition:

$$\frac{f^{(k)}(x) - f^{(k)}(0)}{x - 0} = \frac{1}{x}Q_k\left(\frac{1}{x}\right)\exp\left(-\frac{1}{x^2}\right) = \sum_{s=1}^{3k+1} a_s \frac{\exp(-\frac{1}{x^2})}{x^s}$$

which we must prove tends to zero as $x \to 0$; if we change the variable to $t = \frac{1}{x}$ then we have a finite sum of terms like $t^s \exp -t^2$ which we know tend to zero as $|t|$ tends to infinity.

So for this function f the series $\sum \frac{f^{(k)}(0}{k!}x^k = 0$ so converges to 0 at every x. But the error term $E_n(x)$ is the same for all n (it equals $f(x)$) and so does not tend to 0 at any point except 0.

Note that we can add this function to $\exp x$ and $\sin x$ and so on, and get functions with the same set of derivatives at 0 as these functions, so that they will have the same Taylor polynomials—but are different functions.

Remark 15.1. *Functions defined and differentiable on $\mathbb{C}$ are very different: for them, our naive intuition is a good guide—but that is next year's Analysis course.*

15.3 Taylor's Theorem with Lagrange Remainder

We now concentrate on the Taylor polynomial and investigate its difference from the function.

Theorem 15.2 (Taylor's Theorem)**.** *Let $f : [a,b] \to \mathbb{R}$. Suppose that for some $n \geqslant 1$ we have that f, $f', f'', \ldots, f^{(n-1)}$ exist and are continuous on $[a,b]$ and that $f^{(n)}$ exists on (a,b).*

Then there is a number $\xi \in (a,b)$ such that

$$f(b) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{f^{(n)}(\xi)}{n!} (b-a)^n \qquad (T_n).$$

Note 15.3. *Recall that at the end points a and b 'differentiable' means 'left- (or right-) differentiable.*

Note 15.4. *The term $\frac{f^{(n)}(\xi)}{n!} (b-a)^n$ is called Lagrange's form of the remainder. Note that the crucial parameter ξ, may depend on (i) the function f; (ii) the degree n; (iii) the end points a and b.*[13]

Note 15.5. *If we set $b - a = h$, then Taylor's theorem may be stated as*

$$f(a+h) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} h^k + \frac{f^{(n)}(a+\theta h)}{n!} h^n$$

where θ is some number between 0 and 1.

Proof. We use the method of "varying a constant": Consider $F : [a,b] \to \mathbb{R}$

$$F(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(x)}{k!} (b-x)^k .$$

F is clearly continuous on $[a,b]$ and on (a,b) we have that

$$\begin{aligned} F'(x) &= \sum_{k=0}^{n-1} \frac{f^{(k+1)}(x)}{k!} (b-x)^k + \sum_{k=0}^{n-1} \frac{f^{(k)}(x)}{k!} (-1) k (b-x)^{k-1} \\ &= \sum_{k=1}^{n} \frac{f^{(k)}(x)}{(k-1)!} (b-x)^{k-1} - \sum_{k=1}^{n-1} \frac{f^{(k)}(x)}{(k-1)!} (b-x)^{k-1} = \frac{f^{(n)}(x)}{(n-1)!} (b-x)^{n-1} . \end{aligned}$$

Note also that $F(a) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k$ and $F(b) = f(b)$.

[13] When applying Taylor's Theorem to different functions (perhaps as similar as $f(x)$ and $f(-x)$) or different ranges (perhaps as similar as $[0,b]$ and $[-b,0]$) it is essential to use a different letter for each ξ that is introduced.

Let $G(x)$ be continuous on $[a,b]$ and differentiable on (a,b). We use Cauchy's Mean Value Theorem on this pair of functions to see that there exists a $\xi \in (a,b)$ such that

$$\frac{F(a)-F(b)}{G(a)-G(b)} = \frac{F'(\xi)}{G'(\xi)}.$$

That is

$$\frac{\sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k - f(b)}{G(a)-G(b)} = \frac{\frac{f^{(n)}(\xi)}{(n-1)!} (b-\xi)^{n-1}}{G'(\xi)}. \qquad (*)$$

But if we take

$$G(x) := (b-x)^n,$$

which is clearly continuous on $[a,b]$ and differentiable on (a,b) with derivative $-n(b-x)^{n-1} < 0$, then (*) simplifies at once to

$$f(b) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{f^{(n)}(\xi)}{n!} (b-a)^n .$$

□

Note that, under the conditions given above, we can replace b in (T_n) by any $x \in [a,b]$.

We have proved the strongest theorem we could. But often we know a bit more, and can get, for example this symmetric version:

Corollary 15.6 (Taylor's Theorem)**.** *Let $f : (a-\delta, a+\delta) \to \mathbb{R}$ for some $\delta > 0$. Suppose that for some $n \geqslant 1$ we have that $f', f'', \ldots, f^{(n)}$ exist. Let $x \in (a-\delta, a+\delta)$. Then there is a number ξ between a and x such that*

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (x-a)^k + \frac{f^{(n)}(\xi)}{n!} (x-a)^n .$$

Proof. If $x > a$ then this is just the Taylor Theorem we have proved. If $x < a$ we just use the Taylor Theorem we have proved on the function $f(-x)$. If $x = a$ then take $\xi = a$ (we define $0^0 := 1$ - see also Section 16.3). □

15.4 Other forms of the remainder

In the proof of Taylor's Theorem we may use *any* function G which is continuous in $[a,b]$, differentiable in (a,b), and such that $G' \neq 0$. Then we will have a $\xi \in (a,b)$ such that

$$f(b) = P_{n-1}(b) + \frac{f^{(n)}(\xi)}{(n-1)!} (b-\xi)^{n-1} \frac{G(b)-G(a)}{G'(\xi)}.$$

By choosing different functions G, you may prove Taylor's Theorem with the remainder of different forms. For example, if we choose $G(x) = x - a$, then $\frac{G(b)-G(a)}{G'(\xi)} = b - a$. Thus

$$f(b) = P_{n-1}(b) + \frac{f^{(n)}(\xi)}{(n-1)!}(b-a)(b-\xi)^{n-1} \quad \text{for some } \xi \in (a,b).$$

Exercise 15.1. *Try $G(x) = (x-a)^m$ for a power $m \geqslant 1$ to see what kind of Taylor's formula you can get.*

15.5 The error estimate

Taylor's Theorem also provides us with an explicit estimate of the difference between $f(x)$ and its n-Taylor approximation $\sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k$:

Corollary 15.7. *Let $f : [a,b] \to \mathbb{R}$ satisfy the conditions in Taylor's Theorem, and let $E_n := \frac{|b-a|^n}{n!} \sup_{\xi \in (a,b)} |f^{(n)}(\xi)|$. Then*

$$\left| f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k \right| \leqslant E_n \quad \textit{for all } x \in [a,b].$$

Of course this may not be useful, as the supremum may be infinite. If however in a given situation we know a bit more—for example, that $f^{(n)}$ is differentiable on $[a,b]$ then we can use standard calculus to evaluate E_n.

15.6 Example: the function $\log(1+x)$

By way of an example we prove the following:

Proposition 15.8. *We have*

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} \quad \textit{for all } x \in (-1,1].$$

Note that this is the best result we can get as the radius of convergence of the series is by the ratio test equal to 1 and at the other end point $x = -1$ the series is the notoriously divergent Harmonic Series $\sum \frac{1}{n}$.

But we will prove equality on all of $(-1,1]$, in particular that $\log 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$.

Proof. Consider $f(x) = \log(1+x)$. We have already proved that on $(-1,\infty)$ the function f is differentiable with $f'(x) = \frac{1}{1+x}$; and so, by induction we have $f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n}$ for all $n \geqslant 1$.

Hence, by Taylor's Theorem (the symmetric version)

$$\log(1+x) - \sum_{k=1}^{n-1} \frac{(-1)^{k-1}x^k}{k} = (-1)^{n-1}\frac{1}{n}\left(\frac{x}{1+\xi_n}\right)^n$$

for some ξ_n between 0 and x.

To get our result it would be enough to show that

$$\left|\frac{x}{1+\xi_n}\right| \leqslant 1$$

for every n and $x \in (-1,1]$.

For $x > 0$ this is no problem, $0 < \xi_n < 1$ and so $1+\xi_n > 1$; hence $\frac{x}{1+\xi_n} < x \leqslant 1$.

For negative x it is not so easy; the nearer x is to -1 the nearer $1+\xi_n$ may get to 0. However, if $x \geqslant -\frac{1}{2}$ we have

$$-\tfrac{1}{2} \leqslant x \leqslant \xi_n \leqslant 0$$

and so

$$\tfrac{1}{2} \leqslant 1+\xi_n \leqslant 1$$

which implies

$$2x \leqslant \frac{x}{1+\xi_n} \leqslant x.$$

Now $2x \geqslant -1$ and $x \leqslant 1$ so we have

$$\left|\frac{x}{1+\xi_n}\right| \leqslant 1$$

as required.

That is, the functions $\log(1+x)$ and $\sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k}{k}$ are equal on $[-\frac{1}{2}, 1]$.

What about $(-1, -\frac{1}{2})$? We must use a very different argument.

Consider the functions $f(x) = \log(1+x)$ and $g(x) = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k}{k}$ on $(-1,1)$. Both are differentiable there; we have proved $f'(x) = \frac{1}{1+x}$, and by the theorem on term-by-term differentiation of power series

$$g'(x) = \sum_{k=1}^{\infty}(-1)^{k-1}k\frac{x^{k-1}}{k} = \sum_{k=1}^{\infty}(-1)^{k-1}x^{k-1} = \frac{1}{1+x}.$$

Hence $f'(x) - g'(x) = 0$, so by the Constancy Theorem,

$$f(x) - g(x) = f(0) - g(0) = 0.$$

That is, on the whole of $(-1,1]$ we have the required series expansion.

□

Remark 15.9. *The last part has actually proved the result for $x \in (-1,1)$. It is only at $x = 1$ that we have to prove that the error tends to zero.*

16 The Binomial Theorem

In this section we use many of the theorems we have proved about uniform convergence and continuity, power series, monotonicity as well as Taylor's Theorem. As well as proving an important result we are showing off the techniques we now have available to us.

16.1 Motivation and Preliminary Algebra

By simple induction we can prove that for any natural number n (including 0) we have for all real or complex x that

$$(1+x)^n = \sum_{k=0}^{k=n} \binom{n}{k} x^k;$$

where the coefficient $\binom{n}{k}$ of x^k can be proved to be

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k \cdot (k-1) \cdot \dots \cdot 1}.$$

We have also seen in our work on sequences and series that

$$(1+x)^{-1} = \sum_{k=0}^{\infty} (-1)^k x^k \qquad \text{for all } \ |x| < 1$$

and here the coefficient of x^k can be written as

$$(-1)^k = \frac{(-1) \cdot (-2) \cdot \dots \cdot (-k)}{k \cdot (k-1) \cdot \dots \cdot 1};$$

and we can prove by induction (for example using differentiation term by term) that for all natural numbers $n \geqslant 1$ we have that

$$(1+x)^{-n} = \sum_{k=0}^{\infty} \frac{(-n) \cdot (-n-1) \cdot \dots \cdot (-n-k+1))}{k \cdot (k-1) \cdot \dots \cdot 1} x^k \qquad \text{for all } \ |x| < 1,$$

so the binomial theorem above holds for all integers n.

In this section we are going to generalise these—in the case of some real values of x—to all values of n, not just integers. Note that this is altogether deeper: $(1+x)^p$ is defined for non-integral p, and for (real) $x > -1$, to be the function $\exp(p \log(1+x))$.

Definition 16.1. *For all $p \in \mathbb{R}$ and all $k \in \mathbb{N}$ we extend the definition of* ***binomial coefficient*** *as follows:*

$$\binom{p}{0} := 1; \qquad \textit{and} \qquad \binom{p}{k} := \frac{p(p-1)\dots(p-k+1)}{k!}.$$

We now make sure that the key properties of binomial coefficients are still true in this more general setting.

Lemma 16.1.
$$k\binom{p}{k} = p\binom{p-1}{k-1}, \qquad \textit{for all } k \geqslant 1.$$

Proof. If $k = 1$ then by the definition we must see $1 \cdot \frac{p}{1} = p \cdot 1$ which is clear. Otherwise
$$k\binom{p}{k} = k\frac{p(p-1)\dots(p-k+1)}{k!} = p\frac{(p-1)\dots(p-k+1)}{(k-1)!} = p\binom{p-1}{k-1}.$$
□

Lemma 16.2.
$$\binom{p}{k} + \binom{p}{k-1} = \binom{p+1}{k}, \qquad \textit{for all } k \geqslant 1.$$

Proof. When $k = 1$ we must prove $\frac{p}{1} + 1 = \frac{p+1}{1}$ which is clear. Otherwise
$$\begin{aligned}\binom{p}{k} + \binom{p}{k-1} &= \frac{p(p-1)\dots(p-k+1)}{k!} + \frac{p(p-1)\dots(p-k+2)}{(k-1)!} \\ &= \frac{p(p-1)\dots(p-k+2)}{k!}\left[(p-k+1)+k\right] \\ &= \frac{(p+1)p(p-1)\dots(p-k+2)}{k!} \\ &= \binom{p+1}{k}.\end{aligned}$$
□

16.2 The Real Binomial Theorem

Theorem 16.3 (The Binomial Expansion)**.** *Let p be a real number. Then*
$$(1+x)^p = \sum_{k=0}^{\infty}\binom{p}{k}x^k \qquad \textit{for all } \ |x| < 1 \qquad\qquad (B).$$

Note that the coefficients are all non-zero provided p is not a natural number or zero; as we have a proof of the expansion in that case we may assume that $p \notin \mathbb{N} \cup \{0\}$.

Lemma 16.4. *The function f defined on $(-1, 1)$ by $f(x) := (1+x)^p$ is differentiable, and satisfies $(1+x)f'(x) = pf(x)$. Also, $f(0) = 1$.*

Proof. The derivative is easily got by the chain rule from the definition of f; it is $f'(x) = p(1+x)^{p-1}$. Multiply by $(1+x)$ and get the required relationship. The value at 0 is clear. $\square$

Lemma 16.5. *The radius of convergence of* $\sum_{k=0}^{\infty} \binom{p}{k} x^k$ *is* $R = 1$.

Proof. Use the ratio test; we have that "$|a_{k+1}x^{k+1}/a_k x^k|$" is

$$\left| \frac{p \cdot (p-1) \cdot \dots \cdot (p-k)}{(k+1) \cdot k \cdot (k-1) \cdot \dots \cdot 1} \cdot \frac{k \cdot (k-1) \cdot \dots \cdot 1}{p \cdot (p-1) \cdot \dots \cdot (p-k+1)} x \right| = \left| \frac{p-k}{k+1} x \right| \to |x|$$

as $k \to \infty$. $\square$

Lemma 16.6. *The function g defined on $(-1,1)$ by $g(x) = \sum_{k=0}^{\infty} \binom{p}{k} x^k$ is differentiable, with derivative satisfying $(1+x)g'(x) = pg(x)$. Also, $g(0) = 1$.*

Proof.

$$\begin{aligned}
(1+x)g'(x) &= (1+x) \sum_{k=0}^{\infty} \binom{p}{k} k x^{k-1}, \text{ differentiation term by term valid for } |x| < 1, \\
&= (1+x) \sum_{k=1}^{\infty} \binom{p}{k} k x^{k-1} \\
&= p(1+x) \sum_{k=1}^{\infty} \binom{p-1}{k-1} x^{k-1} \quad \text{by Lemma 16.1} \\
&= p \left\{ \sum_{k=1}^{\infty} \binom{p-1}{k-1} x^{k-1} + \sum_{k=1}^{\infty} \binom{p-1}{k-1} x^k \right\} \\
&= p \left\{ \sum_{m=0}^{\infty} \binom{p-1}{m} x^m + \sum_{m=1}^{\infty} \binom{p-1}{m-1} x^m \right\} \\
&= p \left\{ 1 + \sum_{m=1}^{\infty} \binom{p-1}{m} x^m + \sum_{m=1}^{\infty} \binom{p-1}{m-1} x^m \right\} \\
&= p \left\{ 1 + \sum_{m=1}^{\infty} \left[\binom{p-1}{m} + \binom{p-1}{m-1} \right] x^m \right\} \\
&= p \left\{ 1 + \sum_{m=1}^{\infty} \binom{p}{m} x^m \right\} \quad \text{by Lemma 16.2} \\
&= p \sum_{m=0}^{\infty} \binom{p}{m} x^m \\
&= pg(x).
\end{aligned}$$

$\square$

Proof of the Binomial Theorem. Consider $\phi(x) = \dfrac{g(x)}{f(x)}$, which is well-defined on $(-1, 1)$ as $f(x) > 0$. By the Quotient Rule we can calculate $\phi'(x)$, and then use the lemmas:

$$\phi'(x) = \frac{f(x)g'(x) - f'(x)g(x)}{f(x)^2} = \frac{p}{1+x}\frac{f(x)g(x) - f(x)g(x)}{f(x)^2} = 0.$$

Hence by the Constancy Theorem, $\phi(x)$ is constant, $\phi(x) = \phi(0) = 1$. This implies that $f(x) = g(x)$ on $(-1, 1)$. $\square$

16.3 The end points: preliminary issue

The existence of these functions and their equality at the end points requires more sophisticated argument. Most of the following will probably be omitted from the lectures but in any case **the following sections should be viewed as illustrations of the way Taylor's Theorem can be exploited, rather than theorems to be learnt.**

The cases $x = 1$ or $x = -1$ need to be considered separately. But there is a difference between these!

For $x = -1$ we have not yet defined $(1 + x)^p$.

For $p \in \mathbb{N}$ we have the usual algebraic definition, so $0^p = 0$. Can we define 0^p sensibly for any other values of p?

For $p > 0$: If $x > -1$ we defined $(1+x)^p := \exp p \log(1+x)$. As $\log(1+x) \to -\infty$ as $x \to -1$, we have $\exp p \log(1 + x) \to 0$ as $x \to -1$. Thus to make $(1 + x)^p$ continuous at $x = -1$ we should define $0^p = 0$. This we now do.

If $p = 0$: How one defines 0^0 depends on the context. (Sometimes $0^0 := 1$ sometimes $0^0 := 0$.) If (B) is to hold for $x = 0$ then we must define $0^0 = 1$. But if we do this , then to preserve the rule of exponents $A^pA^q = A^{p+q}$ we cannot define negative powers; if $p > 0$ then 0^{-p} makes no sense.

So let us extend out definition of $(1 + x)^p$ in this way, in the case when $p > 0$.

But we need to take care.

Lemma 16.7. *If $p > 0$ then the function $(1 + x)^p$ is continuous on $[-1, \infty)$.*

Lemma 16.8. *If $p > 1$ then the function $(1+x)^p$ is differentiable on $[-1, \infty)$ with derivative $p(1 + x)^{p-1}$.*

Proofs. Exercises. $\square$

16.4 The end points: $p \leqslant -1$

Let $p \leqslant -1$. Then as remarked above, the function $(1+x)^p$ is not defined at $x = -1$. Further the expansion does not converge at $x = 1$:

Proposition 16.9. *The series* $\sum_{k=0}^{\infty} \dfrac{p \cdot (p-1) \cdot \cdots \cdot (p-k+1)}{k \cdot (k-1) \cdot \cdots \cdot 1}$ *is divergent.*

Proof. Write $q = -p \geqslant 1$; then the modulus of the k-th term

$$\left| \frac{p \cdot (p-1) \cdot \cdots \cdot (p-k+1)}{k \cdot (k-1) \cdot \cdots \cdot 1} \right| = \left| (-1)^k \frac{q}{1} \cdots \frac{q+s}{s+1} \cdots \frac{q+k-1}{k} \right| \geqslant 1;$$

the terms alternate in sign but as they do not tend to 0 the series diverges. $\square$

16.5 The end points: $-1 < p < 0$

Let $-1 < p < 0$; note that $p + 1 > 0$. Again the function $(1+x)^p$ is not defined at $x = -1$. However, now the expansion converges at $x = 1$:

Proposition 16.10. *The series* $\sum_{k=0}^{\infty} \dfrac{p \cdot (p-1) \cdot \cdots \cdot (p-k+1)}{k \cdot (k-1) \cdot \cdots \cdot 1}$ *is convergent with sum* 2^p.

Proof. We apply Taylor's Theorem to $(1+x)^p$ on the interval $[0,1]$ and find, for each $n \geqslant 1$, a point $\xi_n \in (0,1)$ such that

$$2^p = \sum_{k=0}^{n-1} \frac{p \cdot (p-1) \cdot \cdots \cdot (p-k+1)}{k \cdot (k-1) \cdot \cdots \cdot 1} + E_n$$

where

$$E_n = \frac{p \cdot (p-1) \cdot \cdots \cdot (p-n+1)}{n \cdot (n-1) \cdot \cdots \cdot 1} (1+\xi_n)^{p-n}.$$

We have then that

$$|E_n| \leqslant \left| \frac{p \cdot (p-1) \cdot \cdots \cdot (p-n+1)}{n \cdot (n-1) \cdot \cdots \cdot 1} \right|;$$

and we will have the result if we prove that this tends to 0 as $n \to \infty$. We rewrite $|E_n|$ as

$$\begin{aligned} &\left| \frac{[(p+1)-1] \cdot \cdots \cdot [(p+1)-s] \cdot \cdots \cdot [(p+1)-n]}{1 \cdot 2 \cdot \cdots \cdot n} \right| \\ &= \left(1 - \frac{p+1}{1}\right) \cdot \left(1 - \frac{p+1}{2}\right) \cdots \left(1 - \frac{p+1}{s}\right) \cdots \left(1 - \frac{p+1}{n}\right). \end{aligned}$$

Now $\exp(-x) + x - 1$ has positive derivative on $(0,1)$ so by the MVT we have that

$$\left(1 - \frac{p+1}{s}\right) \leqslant \exp\left(-\frac{p+1}{s}\right)$$

so that

$$|E_n| \leqslant \exp\left(-(p+1)\sum_{s=1}^{n} \frac{1}{s}\right).$$

As the harmonic series diverges and $(p+1) > 0$, we get that $E_n \to 0$ as $n \to \infty$.

$\square$

16.6 The end points: $0 < p$

Let $0 < p$. In this case the expansion is valid at $x = 1$ and $x = -1$.

Proposition 16.11. *The series* $\sum_{k=0}^{\infty} \dfrac{p \cdot (p-1) \cdot \dots \cdot (p-k+1)}{k \cdot (k-1) \cdot \dots \cdot 1}$ *is convergent with sum* 2^p*; and the series* $\sum_{k=0}^{\infty} \dfrac{p \cdot (p-1) \cdot \dots \cdot (p-k+1)}{k \cdot (k-1) \cdot \dots \cdot 1}(-1)^k$ *is convergent with sum* 0.

Proof. The end point $x = +1$ is straightforward; use Taylor's Theorem as before and consider the error estimate

$$E_n = \frac{p \cdot (p-1) \cdot \dots \cdot (p-n+1)}{n \cdot (n-1) \cdot \dots \cdot 1}(1+\xi_n)^{p-n}$$

for some $\xi_n \in (0,1)$. Then

$$|E_n| \leqslant \frac{p}{n} \left| \frac{(p-1) \cdot \dots \cdot (p-n+1)}{1 \cdot 2 \cdot \dots \cdot (n-1)} \right| \frac{2^p}{1^n}.$$

Now $\left| \dfrac{p-s}{s} \right| \leqslant 1$ whenever $2s \geqslant p$; so we get that

$$|E_n| \leqslant \frac{p}{n} \left| \frac{(p-1) \cdot \dots \cdot (p - \left[\frac{p}{2}\right])}{1 \cdot 2 \cdot \dots \cdot (\left[\frac{p}{2}\right])} \right| \frac{2^p}{1^n} \to 0 \quad \text{as} \quad n \to \infty$$

as required. The end point $x = -1$ is more difficult. What we do is prove that the sum converges. Noting that as soon as $k \geqslant p+1$ all the terms have the same sign, we see that this means we have proved that the series is absolutely convergent. Now by the properties of power series $\sum_{k=0}^{\infty} \dfrac{p \cdot (p-1) \cdot \dots \cdot (p-k+1)}{k \cdot (k-1) \cdot \dots \cdot 1} x^k$ is absolutely convergent on $(-1,1)$. In particular we have that the series is absolutely convergent on the closed interval $[-1,0]$. Hence the series is uniformly convergent on that interval; and so the series is continuous on $[-1,0]$. As the series is equal to $(1+x)^p$ on $(-1,0]$ we have by continuity that there is equality at -1 as well.

So we must prove that the series converges. We claim that if we can prove this for any p then we can prove it for $(p+1)$. This is because for all $n \geqslant 2p+2$ we have that $\left| \dfrac{p+1}{p-n+1} \right| \leqslant 1$; this allows us to compare the n-th terms and see that those for $(p+1)$ are smaller in modulus. As both series are ultimately the series of terms of constant sign, the comparison test will yield that convergence for p yields convergence for $(p+1)$. So assume from now that $0 < p < 1$; it will suffice to deal with this case.

The modulus of the n-th term can then be written

$$|u_n| = \frac{p}{n} \left(1 - \frac{p}{1}\right) \dots \left(1 - \frac{p}{s}\right) \dots \left(1 - \frac{p}{n-1}\right)$$

and so, using again $(1-t) \leqslant \exp(-t)$, we have that

$$\begin{aligned}|u_n| \leqslant & \frac{p}{n} \exp\left(-p \sum_{s=1}^{n-1} \frac{1}{s}\right) \\ = & \frac{p}{n} \exp\left(-p\left(\sum_{s=1}^{n-1} \frac{1}{s} - \log n\right)\right) \exp\left(-p(\log n)\right) \\ = & \frac{p}{n} \frac{1}{n^p} \exp\left(-p\left(\sum_{s=1}^{n-1} \frac{1}{s} - \log n\right)\right).\end{aligned}$$

Now we have (Integral Test argument) that

$$\sum_{s=1}^{n-1} \frac{1}{s} - \log n \to \gamma \quad \text{as } n \to \infty \qquad (\gamma \text{ is Euler's constant}).$$

Hence we have a constant C such that

$$|u_n| \leqslant C \frac{1}{n \cdot n^p} \quad \text{for sufficiently large } n,$$

and so, by the Comparison Test, $\sum |u_n|$ converges.[14]

$\square$

[14] $\sum \frac{1}{n^s}$ is convergent for $s > 1$ by the Integral Test.